Semidefinite programming approaches for structured combinatorial optimization problems

Dobre, C.

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

Semidefinite programming approaches for structured combinatorial optimization problems

Semidefinite programming approaches for structured combinatorial optimization problems

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit van Tilburg, op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op vrijdag 18 maart 2011 om 10.15 uur door

CRISTIAN DOBRE

geboren op 25 april 1979 te Braşov, Roemenië.

PROMOTOR: prof. dr. E. de Klerk

COMMISSIELEDEN: prof. dr. ir. E.R. van Dam
prof. dr. ir. W.H. Haemers
prof. dr. ir. D. den Hertog
prof. dr. M. Laurent
dr. R. Sotirov
prof. dr. ir. G.J. Wöginger

THOMAS STIELTJES INSTITUTE
FOR MATHEMATICS

Contents

Acknowledgements

"The important thing is this: to be able at any moment to sacrifice what we are for what we could become."

Charles Du Bos
1882–1939

The social interdependencies of any human being have an important influence on his/her personality, behavior, and quality of work. This thesis, and indeed my life in recent years, has been influenced by many people, some of whom I would like to mention by name.

First of all I would like to thank my parents, **Rodica** and Constantin, for their love and for providing me with an excellent education so I could achieve this important milestone. This thesis is dedicated to my mother who left us too early and who deserved to witness this moment. Special thanks go to my father too, for enduring my absence and for understanding what was best for my future. A special word goes to Florin, whose friendship could be felt 2000 km away from the place where it started; and to Carmen, the best dancing partner I ever met.

I would like to express my gratitude to Bogdan and Remus. They changed my life by putting me in contact with my supervisor while I was on vacation in the Netherlands, at the end of summer 2006. Bogdan was also my friendly housemate for four years and I thank him for understanding my capricious behavior at those times when only close friends know how to react.

My *Alma Mater* is Transilvania University in Brasov, Romania, and I am also deeply indebted to the people there who introduced me to mathematical optimization: **Nicolae Pascu**, Gheorghe Munteanu, Emil Tatomir, Eleonor Ciurea, Moise Cocan, and Ernest Scheiber.

Moving to the Netherlands obliged me to become accustomed with the Dutch language and culture. I remember the great lunches with Leo, Jalal, Renata, and Etienne where I could start practicing my first "Dutch" words, and the dinners with Edwin that improved my Dutch football vocabulary. Many thanks to Robin too, for introducing me to the Dutch lifestyle and being perfect company for many of my social activities.

Throughout this four-year PhD program I shared offices with awesome people: K536 with Renata, K531 with Gijs, and K522 with Miguel. Thank you for the comfort and friendship you offered me while we worked.

The text you are now reading was written in LaTeX. Often one must spend hours to achieve a fancy style or simply to figure out a small mistake that destroys the formatting of the document. I was spared such moments because of the much-appreciated help of Elekke and Gerwald. The text of my scientific articles was possible also because of my coauthors: Dima, Etienne, and Renata. Thank you for teaching me so much of value.

Any piece of work has to be evaluated. The heavy text that follows in this manuscript was carefully read and commented on by my PhD committee, to whom I would like to express my appreciation.

Finally, I want to thank my supervisor, Etienne. As well as being my informal boss he is a great person, who understands how any employee would want to be supervised. He was an excellent guide to the nuances of life in academia and life in general.

Cristian Dobre

Tilburg, December 2010.

List of notation

$\mathbb{R}^n$	:	n-dimensional Euclidean vector space;
$\mathbb{R}^n_+$	:	nonnegative orthant of $\mathbb{R}^{\mathrm{n}}$;
$\mathbb{R}^{m\times n}$	:	space of $m \times n$ real matrices;
$\mathbb{C}^{m\times n}$	:	space of $m \times n$ complex matrices;
A^T	:	transpose of $A \in \mathbb{R}^{m\times n}$;
A_{ij}	:	ijth entry of $A \in \mathbb{R}^{m\times n}$;
$\mathbb{S}^{k\times k}$	=	$\{A \mid A \in \mathbb{R}^{k\times k}, A = A^T\}$: set of *symmetric* matrices;
$A \in \mathbb{C}^{n\times n}$	:	decomposed as $A = \mathrm{Re}(A) + \sqrt{-1}\mathrm{Im}(A)$;
$\mathrm{Re}(A) \in \mathbb{R}^{n\times n}$	:	real part of $A \in \mathbb{C}^{n\times n}$;
$\mathrm{Im}(A) \in \mathbb{R}^{n\times n}$	:	imaginary part of $A \in \mathbb{C}^{n\times n}$;
A^*	=	$\mathrm{Re}(A)^T - \sqrt{-1}\mathrm{Im}(A)^T$ denotes the conjugate transpose;
$\mathbb{H}^{k\times k}$	=	$\{A \mid A \in \mathbb{C}^{k\times k}, A = A^*\}$: set of *Hermitian* matrices;
$A \succeq 0$ ($A \succ 0$)	:	A is Hermitian/symmetric positive semidefinite (positive definite);
$A \preceq 0$ ($A \prec 0$)	:	A is Hermitian/symmetric negative semidefinite (negative definite);
$\mathbb{S}^{k\times k}_+$	=	$\{A \mid A \in \mathbb{S}^{k\times k},\ A \succeq 0\}$;
$\lambda_i(A)$	:	ith largest eigenvalue of matrix A;
$\mathrm{trace}(A)$	=	$\sum_i A_{ii} = \sum_i \lambda_i(A)$;
$\det(A)$	=	$\Pi_i \lambda_i(A)$ (determinant of A);
$\|A\|^2$	=	$\mathrm{trace}(AA^T) = \sum_i \sum_j A_{ij}^2$ (Frobenius norm)
	=	$\sum_i \lambda_i^2(A)$ if $A \in \mathbb{S}^{k\times k}$;
$\langle A, B\rangle$	=	$\mathrm{trace}(AB^T)$;
$A^{\frac{1}{2}}$	:	unique symmetric square root factor of $A \succeq 0$;
$A(\alpha, \beta)$	:	submatrix that contains the rows of A indexed by α and the columns indexed by β, for index sets $\alpha, \beta \subset \{1, \ldots, \mathrm{k}\}$;
$A(\alpha)$	=	$A(\alpha, \alpha)$;
$A(i, :)$	:	ith row of matrix A;
I_n	:	identity matrix of order n;
J_n	:	$n \times n$ all-ones matrix;
e	:	all-ones vector;
e_i	:	ith standard basis vector;
E_{ij}	=	$e_i e_j^T$;
$\mathrm{Diag}(a)$	:	diagonal matrix with components of $a \in \mathbb{C}^n$ on the diagonal;
$\mathrm{diag}(A)$	:	vector obtained by extracting the diagonal of $A \in \mathbb{C}^{n\times n}$;
$\mathrm{vec}(A)$	=	$[A_{11}, A_{21}, ..., A_{n1}, A_{12}, A_{22}, ..., A_{nn}]^T$ for $A \in \mathbb{C}^{n\times n}$;
Π_n	:	set of $n \times n$ permutation matrices;
$\mathcal{S}_n$	:	*symmetric* group on n elements
$A \otimes B$	:	block matrix with block ij given by $A_{ij}B$ (*Kronecker* product).

Chapter 1

Introduction

1.1 Semidefinite programming

Semidefinite programming (SDP) may be described as linear programming (LP) with positive semidefinite matrix variables. For given symmetric $n \times n$ matrices $A_0, \ldots, A_m$ and $b \in \mathbb{R}^m$, the standard SDP problem is defined as:

$$\begin{array}{ll} \inf & \langle A_0, X \rangle \\ \text{s.t.} & \langle A_i, X \rangle = b_i \ (i = 1, \ldots m) \\ & X \succeq 0, \end{array}$$

where $X \succeq 0$ means X must be symmetric positive semidefinite, and the inner product is the Euclidean inner product: $\langle A_i, X \rangle := \text{trace}(A_i X)$, for $i = 0, ..., m$.

Semidefinite programming is currently one of the most active areas of research in mathematical programming. The reason for this is twofold. First, applications of SDP may be found in control theory, combinatorics, real algebraic geometry, global optimization, and structural design, to name only a few; see the surveys by Vandenberghe and Boyd (1996) and Todd (2001) for more information. Secondly, the extension of interior point methods from linear programming to SDP in the 1990's by Nesterov and Nemirovski (1994), Alizadeh (1991), and others, allows the solution of SDP problems in polynomial time to any fixed accuracy.

This thesis considers applications of SDP to combinatorial optimization problems such as computing the crossing number of a graph, computing the clique number of a graph, solving a traveling salesman problem, or finding a maximum equipartition, all of which are known to be NP-hard. That is, there exists no polynomial-time algorithm that can solve these problems to optimality, unless P=NP. Therefore, approximating their optimal solution in polynomial time is an important goal. New techniques have been developed in the last thirty years using semidefinite programming approaches. However, the SDPs involved are often very large and the size of the problems that can be solved is still limited.

The applications mentioned above are not the only ones in the literature. One could also mention the work on symmetry in SDP of Vallentin (2009), SDP bounds on error correcting codes (see Gijswijt, Schrijver, and Tanaka (2006), Schrijver (2005), and Laurent (2009)), SDP bounds on kissing numbers (see Bachoc and Vallentin (2008) and Mittelmann and Vallentin (2010)), connections between SDP relaxations for the maximum cut problem and the computation of the stability number (see Laurent, Poljak, and Rendl (1997)), SDP bounds on the chromatic number (see Gvozdenović and Laurent (2008a), Dukanovic and Rendl (2007), and Gvozdenović and Laurent (2008b)), and engineering applications in truss topology design (see Bai, De Klerk, Pasechnik, and Sotirov (2009)).

A recurrent difficulty in applying interior point methods is that it is more difficult to exploit special structure in the data in the SDP case than in the LP case. In particular, sparsity may be readily exploited by interior point methods in LP, but this is not true for SDP. There are currently three types of structure (apart from general sparsity) that may be exploited in SDP:

- chordal structure (i.e., the data matrices of the SDP problem have a common sparsity pattern that is the same as the sparsity pattern of a *chordal graph*, which is a graph that does not contain a cycle of length 4 or more as an induced subgraph), see Section 3 of De Klerk (2010) for details;

- *low rank* (i.e., the data matrices have low rank), see Section 2 of De Klerk (2010) for details;

- *algebraic symmetry* (the key ingredient used in this thesis), see Chapter 2 for details.

In the same vein, we will present in this thesis results on exploiting symmetry in the data of SDP relaxations for structured combinatorial optimization problems such as those described in the next section.

1.2 Relaxations of combinatorial optimization problems

A key observation that connects SDP relaxations to combinatorial optimization problems is the following:

$$x^T Q x = \text{trace}(Qxx^T), \tag{1.1}$$

for a given vector x and matrix Q. Then, if we define $X := xx^T$ we can rewrite the quadratic expression in x as a linear formulation in the new matrix variable X:

$$\text{trace}(Qxx^T) = \text{trace}(QX).$$

Since any symmetric positive semidefinite matrix X has a factorization $X = LL^T$, for some matrix L (see e.g., Section 7.2 in Horn and Johnson (1990)), we can readily see the equivalence

$$X = xx^T \quad \Leftrightarrow \quad \text{rank}(X) = 1 \text{ and } X \succeq 0. \tag{1.2}$$

The rank constraint is a nonconvex hard constraint, so we omit the constraint rank(X) = 1 and thus relax the condition $X = xx^T$ to $X \succeq 0$. This relaxation was first used in optimization by Shor (1987).

A second key observation is that we can view many combinatorial optimization problems as quadratic optimization problems. A simple illustration is the equivalence between $x_i \in \{-1, 1\}$ and $x_i^2 = 1$. Since it is not immediately obvious that we can benefit from this nonconvex problem reformulation, we give another motivation via the maximum cut problem.

1.2.1 Maximum cut problem

Let $G = (V, E)$ be an undirected graph with the weighted adjacency matrix W. The problem is to partition the set V into two subsets S_1 and S_2 such that the sum of the weights of the edges between S_1 and S_2 is maximized. The combinatorial approach is to introduce for each node $i \in V$ a variable $x_i \in \{-1, 1\}$. Node i must be placed in either S_1 or S_2. Hence, we can assign $x_i = 1$ if $i \in S_1$ and $x_i = -1$ if $i \in S_2$. If $|V| = n$, then the MAX-CUT problem takes the form of the binary quadratic optimization problem

$$\begin{aligned} \text{MC} := \max \quad & \sum_{i,j=1}^{n} W_{ij} \left(\frac{1 - x_i x_j}{4} \right) \\ \text{s.t.} \quad & x_i \in \{-1, 1\}, \quad i \in V. \end{aligned} \tag{1.3}$$

This is indeed true, since if $(i, j) \in E$ and i and j are in the same set, then $1 - x_i x_j = 0$ and the weight of the edge is not added. On the other hand, if for example $i \in S_1$ and $j \in S_2$ then $1 - x_i x_j = 2$. Since we sum over all edges and every edge appears twice in the sum, we need the factor of $\frac{1}{4}$ to get the actual weight of the cut.

Let us now define the vector $x \in \{-1, 1\}^n$ having as entries the variables x_i $(i = 1, ..., n)$ from (1.3). Further, define the matrix $X = xx^T$. Notice that $X_{ij} = x_i x_j$, for any $i, j = 1, ..., n$, hence $X_{ii} = x_i^2 = 1$ (i.e., diag(X)=e). Using (1.2), we can rewrite (1.3) as the equivalent formulation (having X as a variable):

$$\begin{aligned} \text{MC} = \max \quad & \frac{1}{4} \text{trace}(W(J - X)) \\ \text{s.t.} \quad & \text{diag(X)} = \text{e}, \\ & X \succeq 0, \\ & \text{rank}(X) = 1. \end{aligned} \tag{1.4}$$

If A is the adjacency matrix of the graph and $L := Diag(Ae) - A$ denotes the Laplacian matrix of the graph G, we can easily prove that $\text{trace}(W(J - X)) = \text{trace}(LX)$ (see Section 5.2.2).

We obtain the SDP relaxation of the MAX-CUT problem by deleting the rank constraint, to obtain

$$\begin{aligned} \text{SDP}_{\text{MC}} = \max \quad & \frac{1}{4} \text{trace}(LX) \\ \text{s.t.} \quad & \text{diag(X)} = \text{e}, \\ & X \succeq 0. \end{aligned} \tag{1.5}$$

The convex hull of the set $\{xx^T \mid x \in \{-1, 1\}^n\}$ is therefore approximated by the convex *elliptope*:

$$\mathcal{E} := \{X \in \mathbb{R}^{n \times n} \mid \text{diag}(X) = e, \ X \succeq 0\}.$$

With respect to the quality of bound (1.5), Goemans and Williamson (1995) proved that $\text{SDP}_{\text{MC}} \leq 1.138\text{MC}$. This proof is based on the fact that

$$\frac{2}{\pi}\arcsin(\mathcal{E}) \subset \text{conv}\{xx^T \mid x \in \{-1,1\}^n\} \subset \mathcal{E},$$

where the arcsin function is applied entry-wise. Moreover, they derived a randomized algorithm that provides a cut with expected value greater than 0.878MC. In practice this method performs well: the solutions (i.e., cuts) obtained are closer to optimality than is predicted by the theory.

This Shor-type relaxation of quadratically reformulated combinatorial optimization problems has become a powerful theoretical and computational tool, as we will see in the examples presented in this section.

Maximum bisection

If we additionally require the sets S_1 and S_2 to have equal cardinality we obtain the maximum bisection problem. Obviously this is a particular case of the MAX-CUT problem and adding an appropriate constraint that characterizes the equality of $|S_1|$ and $|S_2|$ yields another SDP relaxation, due to Frieze and Jerrum (1997). Equal cardinality of the sets S_1 and S_2 requires that $x \in \{-1,1\}^n$ has an equal number of 1 and -1 entries. Therefore, the sum of the elements of each row in matrix X should be zero. We have

$$\begin{aligned} \text{FJ} = \max \quad & \frac{1}{4}\text{trace}(LX) \\ \text{s.t.} \quad & \text{diag(X)} = \text{e}, \\ & Xe = 0, \\ & X \succeq 0. \end{aligned} \tag{1.6}$$

In Chapter 5 of this thesis we propose another SDP relaxation for this problem and conduct theoretical and numerical comparisons.

1.2.2 Quadratic assignment problem

The definition of the quadratic assignment problem (QAP) is as follows: given two sets, P (“facilities”) and L (“locations”), of equal sizes together with a flow function $b : P \times P \mapsto \mathbb{R}$ and a distance function $a : L \times L \mapsto \mathbb{R}$, the problem is to find a permutation $\pi : P \mapsto L$ ("assignment") that minimizes the function

$$\sum_{i,j \in P} b(i,j) a(\pi(i), \pi(j)).$$

If we see the flow and distance functions as matrices over the reals we can consider the QAP to be the problem of minimizing

$$\sum_{i,j \in P} b_{ij} a_{\pi(i)\pi(j)}. \tag{1.7}$$

This formulation was introduced by Koopmans and Beckmann (1957). Assuming $|P| = |L| = n$, we can represent any permutation π via a permutation matrix X as follows:

$$X_{ij} = \begin{cases} 1 & \text{if } \pi(i) = j \\ 0 & \text{otherwise.} \end{cases} \qquad (i, j = 1, \ldots, n)$$

If we denote by Π_n the set of all permutation matrices, we have the trace formulation of QAP:

$$\min_{X \in \Pi_n} \mathrm{trace}(BXAX^T). \tag{1.8}$$

It is well known that the QAP contains the traveling salesman problem (TSP) as a particular case, when taking for example

$$B = \begin{pmatrix} 0 & 1 & 0 & . & . & 1 \\ 1 & 0 & 1 & 0 & . & . \\ . & 1 & 0 & 1 & 0 & . \\ . & . & . & . & . & . \\ 0 & . & . & . & 0 & 1 \\ 1 & 0 & . & . & 1 & 0 \end{pmatrix}, \tag{1.9}$$

and $A = \frac{1}{2} D$, where D is the distance matrix between the nodes. Therefore, the QAP is an NP-hard problem. QAPs of size $n \geq 25$ are still considered to be difficult, so *branch and bound* algorithms (see Anstreicher (2003)) are used to solve them. In turn these algorithms depend on the quality of the lower bounds computed for the QAP.

Based on the same equivalence from (1.2), Zhao, Karisch, Rendl, and Wolkowicz (1998) and Rendl and Sotirov (2007) lifted the problem from $\mathbb{R}^{n \times n}$ to the positive semidefinite cone of dimension $n^2 + 1$, by considering the matrix variable $Y :=$

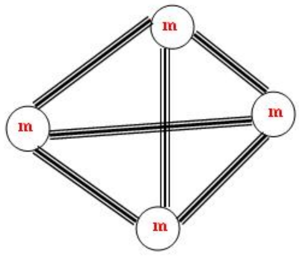

Figure 1.1: A schematic illustration of $K_{m,m,m,m}$

$\begin{pmatrix} 1 \\ \mathrm{vec}(X) \end{pmatrix} \begin{pmatrix} 1 \\ \mathrm{vec}(X) \end{pmatrix}^T$, where X is the permutation matrix from (1.8). Again, $Y \succeq 0$, rank(Y)=1 and the relaxations are obtained by ignoring the nonconvex rank constraint.

Continuing in the same vein, bounds for the QAP were obtained by Povh and Rendl (2009), this time lifting the variable in the cone of n^2-dimensional semidefinite matrices, by setting $Y := \mathrm{vec}(X)\mathrm{vec}(X)^T$. These bounds are equivalent to the bounds of Rendl and Sotirov (2007). Later, improved bounds were obtained by De Klerk and Sotirov (2010b) for certain QAP instances where one may fix one facility to one location without loss of generality.

We have already seen that the TSP is a special case of the QAP. The *maximum k-section* problem is also a special case of the QAP. By maximum k-section we understand partitioning the vertices of a graph into k sets with equal cardinalities such that the sum of the edges between the sets is maximized.

To formulate the maximum k-section problem as a QAP, consider the adjacency matrix of the complete multipartite graph $K_{m,\dots,m}$, see Fig. 1.1 (with any fixed labeling of the vertices), where $n = km$, e.g.,

$$B := (J_k - I_k) \otimes J_m \equiv \begin{pmatrix} 0_m & J_m & \dots & J_m \\ J_m & 0_m & \ddots & \vdots \\ \vdots & \ddots & \ddots & J_m \\ J_m & \cdots & J_m & 0_m \end{pmatrix} \in \mathbb{S}^{km \times km}. \tag{1.10}$$

If X is a permutation matrix that defines a relabeling of the vertices, then the adjacency matrix after relabeling is $X^T B X$.

The QAP reformulation of max k-section on a complete graph with vertex set V

($|V| = km$) and matrix of edge weights W is therefore given by

$$\frac{1}{2} \max_{X \in \Pi_{|V|}} \operatorname{trace}(W X^T B X). \tag{1.11}$$

In the case where $k = 2$ we obtain the maximum bisection problem for which we have already seen relaxation (1.6).

As a consequence, new bounds on the QAP also offer new bounds for the TSP and maximum k-section problems. However, the sizes of these relaxations are considerable. The good news is that they do not have to be solved in this form, for these particular cases.

Exploiting the special structure of the matrix in (1.9), De Klerk, Pasechnik, and Sotirov (2008) have developed a new SDP relaxation for the TSP. Further, in Chapter 4 of this thesis we propose a new relaxation of a special case of the TSP, where the matrix of distances is symmetric and circulant. This special structure actually leads to a linear programming bound. By exploiting the special structure of the matrix in (1.10), we derive in Chapter 5 of this thesis a new relaxation for the maximum k-section problem and compare it to that given in (1.6) (for $k = 2$).

1.2.3 Crossing number of complete bipartite graphs

The crossing number $\operatorname{cr}(G)$ of a graph G is the minimum number of intersections of edges in a drawing of G in the plane. Paul Turán raised the problem of computing the crossing number of a complete bipartite graph $K_{r,s}$; see Turán (1977). The crossing number of the complete bipartite graph is known only in a few special cases (such as $\min\{r, s\} \leq 6$), and it is therefore interesting to obtain lower bounds on $\operatorname{cr}(K_{r,s})$. There is a well-known upper bound on $\operatorname{cr}(K_{r,s})$ via a drawing that is conjectured to be tight. This drawing of $K_{4,5}$ with 8 crossings is presented in Fig. 1.2.

De Klerk, Maharry, Pasechnik, Richter, and Salazar (2006) showed that we can obtain a lower bound on $\operatorname{cr}(K_{r,s})$ via the optimal value of a suitable SDP, namely

$$\operatorname{cr}(K_{r,s}) \geq \frac{s}{2}\left(s \min_{X \geq 0,\, X \succeq 0} \{\operatorname{trace}(MX) \;\mid\; \operatorname{trace}(JX) = 1\} - \left\lfloor \frac{r}{2} \right\rfloor \left\lfloor \frac{r-1}{2} \right\rfloor \right),$$

where M is a certain (given) matrix of order $n = (r-1)!$, and J is the all-ones matrix of the same size. The rows and columns of M are indexed by all the cyclic orderings of r elements. (The cyclic orderings are given by the equivalence classes of orderings that are equal modulo a cyclic permutation.) Therefore, we have $\frac{r!}{r}$ cyclic orderings

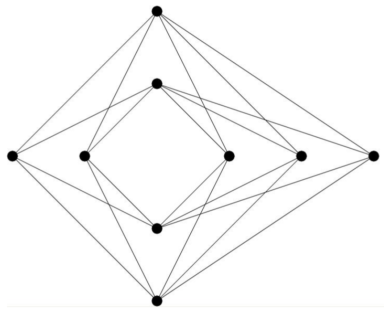

Figure 1.2: Drawing of $K_{4,5}$

that we denote $u_1, ..., u_{(r-1)!}$. The entries M_{ij} are given by the distance between cyclic orderings u_i and u_j. (This distance is given by the number of neighbor swaps needed to go from one ordering to another; for example the distance between 123 and 213 is one.)

De Klerk, Maharry, Pasechnik, Richter, and Salazar (2006) solved the SDP for $r = 7$ using partial symmetry reduction and obtained the bound

$$\mathrm{cr}(K_{7,s}) \geq 2.1796s^2 - 4.5s.$$

Later, De Klerk, Pasechnik, and Schrijver (2007) solved the SDP for $r = 9$ using representation theory and obtained the bound

$$\mathrm{cr}(K_{9,s}) \geq 3.8676063s^2 - 8s.$$

However, when solving the underlying SDP for $r = 9$, the solution time reported by De Klerk, Pasechnik, and Schrijver (2007) was 7 days of wall-clock time on an SGI Altix supercomputer. Using the *numerical symmetry reduction* presented in Chapter 3 of this thesis we can reduce the time to about 24 minutes on a Pentium IV PC, including the time for preprocessing the data.

1.2.4 Stability number, chromatic number, Lovász ϑ, and ϑ'

Let $G = (V, E)$ be an undirected graph without loops. A subset $S \subseteq V$ is called a *stable set* of G if the induced subgraph on S contains no edges. The *maximum stable set problem* is to find a stable set of maximum cardinality. The *stability number*, denoted $\alpha(G)$, is defined as the cardinality of a maximum stable set. The *chromatic number* of G is the minimum number of colors required to color the vertices of G

such that no two adjacent vertices receive the same color. The chromatic number is denoted $\chi(G)$.

The ϑ number, introduced by Lovász (1979), may be defined as the optimal value of the following semidefinite problem:

$$\begin{aligned} \vartheta(G) := \max \quad & \operatorname{trace}(JX) \\ \text{s.t.} \quad & X_{ij} = 0,\ (i,j) \in E \ (i \neq j), \\ & \operatorname{trace}(X) = 1, \\ & X \succeq 0. \end{aligned}$$

One of the best-known properties of the ϑ number is the so-called *sandwich theorem*.

Theorem 1.2.1 (Lovász (1979))**.** *Let $G = (V, E)$ be a graph with stability number $\alpha(G)$, and let $\overline{G}$ be its complementary graph with chromatic number $\chi(\overline{G})$. Then*

$$\alpha(G) \leq \vartheta(G) \leq \chi(\overline{G}).$$

Thus, we can see $\vartheta(G)$ as an SDP upper bound for $\alpha(G)$; but also as an SDP lower bound for $\chi(\overline{G})$.

The ϑ'-number of a graph was introduced by McEliece, Rodemich, and Rumsey (1978) as a strengthening of the Lovász (1979) ϑ-number upper bound on the co-clique number of a graph. Independently, the ϑ'-number was studied in detail for Hamming graphs by Schrijver (1979). The ϑ' number may be defined as the optimal value of the following semidefinite problem:

$$\begin{aligned} \vartheta'(G) := \max \quad & \operatorname{trace}(JX) \\ \text{s.t.} \quad & \operatorname{trace}((A+I)X) = 1, \\ & X \geq 0, \\ & X \succeq 0, \end{aligned}$$

where A is the adjacency matrix of the graph G. This bound is obtained by adding the nonnegativity constraint $X \geq 0$.

Chapter 3 of this thesis uses this example together with the crossing number of bipartite graphs to illustrate a *numerical reduction* technique to reduce the size of the data in semidefinite programs that exhibit *algebraic symmetry*.

1.3 Overview and contribution of this thesis

This thesis is organized into five chapters. Chapter 1 briefly introduced the research questions and Chapter 2 presents the preliminary material necessary for a self-contained thesis. Chapters 3 through 5 give a detailed treatment of the research questions, and they are based on the following three research papers, respectively:

- Klerk, E. de, C. Dobre, and D.V. Pasechnik (2010). Numerical block diagonalization of matrix $*$-algebras with application to semidefinite programming. *Mathematical Programming, Series B*. To appear.

- Klerk, E. de, and C. Dobre (2009). A comparison of lower bounds for the symmetric circulant traveling salesman problem. *Preprint. Tilburg University.* Submitted for publication.

- Klerk, E. de, D.V. Pasechnik, R. Sotirov, and C. Dobre (2010). On semidefinite programming relaxations of maximum k-section. *Preprint. Tilburg University.* Submitted for publication.

The concept of algebraic symmetry is presented in detail in Section 2.5 of Chapter 2, based on the paper of De Klerk, Dobre, and Pasechnik (2010). Some basic ingredients including introductory notes on semidefinite matrices and matrix $*$-algebras are presented in Chapter 2. Examples of matrix $*$-algebras are presented in more detail, since they are key to exploiting symmetry in the SDPs encountered in this thesis. In addition to the preliminary material, Chapter 2 presents research contributions: Section 2.3.3 proves a special structure of the canonical Wedderburn decomposition of the regular *-representation of a matrix $*$-algebra. These results were stated without proof in De Klerk, Dobre, and Pasechnik (2010).

Chapter 3 is in its entirety a contribution to the existing literature. Here, a numerical technique to block diagonalize matrix $*$-algebras is presented. This result is an alternative to the approach by Murota, Kanno, Kojima, and Kojima (2010), and is useful in particular when the initial data set is too large to be handled by the method of Murota *et al.* One important difference between the two methods lies in the underlying $*$-algebra. Whereas Murota, Kanno, Kojima, and Kojima (2010) utilize $*$-algebras over the reals, the technique in Chapter 3 deals with $*$-algebras over the complex numbers. The method we propose is founded on the theorem by Wedderburn (1907), and it is accomplished in two phases. The decomposition in the first phase is carried out not in a given $*$-algebra but in the center of its regular *-representation, which has nice properties that allow numerical computations. Examples on computing bounds for the crossing number of a graph and bounds for the ϑ' number of a graph confirm the relevance of our approach.

In Chapter 4 the attention shifts to deriving a new SDP relaxation for the symmetric circulant traveling salesman problem (SCTSP). By exploiting specific structural and algebraic properties of symmetric circulant matrices we show that in the case of the SCTSP the SDP-based bound from De Klerk, Pasechnik, and Sotirov (2008) can be computed by simply solving a linear programming problem of a size that is polynomial in the size of the input. We perform theoretical and numerical comparisons with the existing bounds in the literature. All the results in this chapter are new and form the content of the paper by De Klerk and Dobre (2009).

Chapter 5 builds on a general framework for exploiting symmetry in a semidefinite relaxation of the QAP due to De Klerk and Sotirov (2010b). First, using the isomorphism from Section 2.2.4 one can reduce the dimension of the initial QAP relaxation from, say n^2, to roughly $2n$. Then, this relaxation is shown to be as least as good as the SDP relaxation due to Karisch and Rendl (1998). All these results are new and appeared in the paper by De Klerk, Pasechnik, Sotirov, and Dobre (2010). Chapter 5 also contains a proof of the equivalence between the bound due to Karisch and Rendl (1998) and the more general QAP bound due to Povh and Rendl (2009), when the latter bound is adapted for the special case of maximum k-section.

Chapter 2

Preliminaries

The aim of this chapter is to present the definitions and basic facts necessary for a self-contained thesis. In the first section we present a fundamental result on matrix $\mathbb{C}*$-algebras, preceded by the necessary definitions: roughly speaking, any matrix $\mathbb{C}*$-algebra can be decomposed as a direct sum of full matrix $\mathbb{C}*$-algebras. The second section is dedicated to examples of matrix $*$-algebras over the reals, since they will feature again in this thesis. The regular *-representation used by De Klerk, Pasechnik, and Schrijver (2007) to reduce semidefinite programs is presented in Section 2.3. Basic results on positive semidefinite matrices and semidefinite programming are grouped in Section 2.4. We end this chapter with some nontrivial results on the symmetry reduction of semidefinite programs that will be used throughout the thesis.

The following properties of the Kronecker product will also be used, see e.g., Graham (1981) (we assume that the dimensions of the matrices appearing in these identities are such that all expressions are well defined):

$$(A \otimes B)(C \otimes D) = AC \otimes BD. \tag{2.1}$$

$$\mathrm{trace}(A \otimes B) = \mathrm{trace}(A)\mathrm{trace}(B). \tag{2.2}$$

$$\exists P \in \Pi_{n^2} \ s.t. \ \forall A, B \in \mathbb{R}^{n\times n}: \quad P(B \otimes A)P^T = A \otimes B. \tag{2.3}$$

Moreover, following the notation already introduced, it can easily be verified that for any column vectors $v, w \in \mathbb{R}^n$:

$$\mathrm{Diag}(\mathrm{vec}(vw^T)) = \mathrm{Diag}(w) \otimes \mathrm{Diag}(v). \tag{2.4}$$

2.1 Basic properties of matrix $*$-algebras

In what follows we give a review of decompositions of matrix $*$-algebras over $\mathbb{C}$, with an emphasis on the constructive (algorithmic) aspects.

Definition 2.1.1. *A set $\mathcal{A} \subseteq \mathbb{C}^{n\times n}$ is called a matrix $*$-algebra over $\mathbb{C}$ (or a matrix $\mathbb{C}*$-algebra) if, for all $X, Y \in \mathcal{A}$:*

- $\alpha X + \beta Y \in \mathcal{A} \quad \forall \alpha, \beta \in \mathbb{C}$;
- $X^* \in \mathcal{A}$;
- $XY \in \mathcal{A}$.

A matrix $\mathbb{C}*$-subalgebra of $\mathcal{A}$ is said to be maximal if it is not contained in any proper $\mathbb{C}*$-subalgebra of $\mathcal{A}$. (Recall that a subset of a set is proper if it is not the empty set or the set itself.)

In applications one often encounters matrix $\mathbb{C}*$-algebras with the following additional structure.

Definition 2.1.2. *Assume that a given set of zero-one $n \times n$ matrices $\{A_1, \ldots, A_d\}$ has the following properties:*

(1) $\sum_{i\in\mathcal{I}} A_i = I$ *for some index set* $\mathcal{I} \subset \{1, \ldots, d\}$ *and* $\sum_{i=1}^{d} A_i = J$;

(2) $A_i^T \in \mathcal{A}$ *for each* i;

(3) $A_iA_j \in span\{A_1, \ldots, A_d\}$ *for all* i, j.

Then $\{A_1, \ldots, A_d\}$ *is called a* coherent configuration.

Thus, a coherent configuration is a basis of zero-one matrices of a (possibly noncommutative) matrix $*$-algebra. Such an algebra is called a *coherent algebra*. Moreover, when the elements of the set $\{A_1, \ldots, A_d\}$ commute and $I \in \{A_1, ..., A_d\}$, the basis of zero-one matrices is called an *association scheme.*

The following results will be useful for developing the theory of matrix $\mathbb{C}*$-algebra decomposition in Chapter 3.

Proposition 2.1.1 (see e.g., Section 1.5 in Godsil (2005))**.** *The elements of a commutative matrix $\mathbb{C}*$-algebra have a common set of orthonormal eigenvectors. These may be viewed as the columns of a unitary matrix Q, i.e., $Q^*Q = I$.*

Proposition 2.1.2 (see e.g., Section 1.5 in Godsil (2005)). *Let $\mathcal{A} \subseteq \mathbb{C}^{n\times n}$ be a commutative matrix $\mathbb{C}*$-algebra of dimension t containing the identity. There exists a basis $E_1, \ldots, E_t$ of $\mathcal{A}$ with the following properties:*

(1) $E_i = E_i^2$ for all i (idempotent);

(2) $\sum_{i=1}^{t} E_i = I$;

(3) $E_i = E_i^*$ for all i (self-adjoint);

(4) $E_iE_j = 0$ if $i \neq j$ (orthogonal).

More information on coherent configurations and related structures may be found in the papers by Higman (1987) and Cameron (2003).

For matrices A_1, A_2, the direct sum is defined as

$$A_1 \oplus A_2 := \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \tag{2.5}$$

and we will denote the iterated direct sum of $A_1, ..., A_n$ by $\bigoplus_{i=1}^{n} A_i$. If all A_i are equal we define:

$$t \odot A := \bigoplus_{i=1}^{t} A.$$

Let $\mathcal{A}$ and $\mathcal{B}$ be two matrix $\mathbb{C}*$-algebras. Then the direct sum of $\mathcal{A}$ and $\mathcal{B}$ is:

$$\mathcal{A} \oplus \mathcal{B} := \{M \oplus M' \mid M \in \mathcal{A}, M' \in \mathcal{B}\}.$$

We say that $\mathcal{A}$ is a zero algebra if $\mathcal{A}$ consists only of the zero matrix.

Definition 2.1.3. *A matrix $\mathbb{C}*$-algebra is called simple if it has no nontrivial ideal. (An ideal of $\mathcal{A}$ is a $*$-subalgebra that is closed under both left and right multiplication by elements of $\mathcal{A}$.)*

Definition 2.1.4. *A matrix $\mathbb{C}*$-algebra is called basic if*

$$\mathcal{A} = t \odot \mathbb{C}^{s\times s} := \{t \odot M \mid M \in \mathbb{C}^{s\times s}\} \tag{2.6}$$

for some integers s, t.

Definition 2.1.5. *Two matrix $\mathbb{C}*$-algebras $\mathcal{A}, \mathcal{B} \subset \mathbb{C}^{n\times n}$ are called equivalent if there exists a unitary matrix $Q \in \mathbb{C}^{n\times n}$ such that*

$$\mathcal{B} = \{Q^* M Q \mid M \in \mathcal{A}\} =: Q^* \mathcal{A} Q.$$

Proposition 2.1.3 (see e.g., Section 2.2 in Gijswijt (2005))**.** *Every matrix $\mathbb{C}*$-algebra $\mathcal{A}$ containing the identity is equivalent to a direct sum of simple matrix $\mathbb{C}*$-algebras.*

Proposition 2.1.4 (see e.g., Section 2.2 in Gijswijt (2005))**.** *Every simple matrix $\mathbb{C}*$-algebra $\mathcal{A}$ containing the identity is equivalent to a basic matrix $\mathbb{C}*$-algebra.*

Propositions 2.1.3 and 2.1.4 imply the so-called fundamental structure theorem for matrix C*-algebras, which is as follows:

Theorem 2.1.5 (Wedderburn (1907))**.** *If $\mathcal{A} \subseteq \mathbb{C}^{n\times n}$ is a matrix $*$-algebra that contains the identity, then there exist a unitary matrix Q and positive integers p and n_i, t_i $(i = 1, \dots, p)$ such that*

$$Q^* \mathcal{A} Q = \oplus_{i=1}^{p} t_i \odot \mathbb{C}^{n_i \times n_i}.$$

Thus, $\dim(\mathcal{A}) = \sum_{i=1}^{p} n_i^2$ and $n = \sum_{i=1}^{p} t_i n_i$.

If the identity does not belong to $\mathcal{A}$, then in view of Definition 2.1.5, each matrix $*$-algebra over $\mathbb{C}$ is equivalent to a direct sum of basic algebras and possibly a zero algebra. A detailed proof of this result is given e.g., in the thesis of Gijswijt (2005) (Theorem 1 there). The proof is constructive and forms the basis for numerical procedures that obtain the decomposition into basic algebras.

Based on the theorem, we define the $*$-isomorphism:

$$\phi : \mathcal{A} \mapsto \oplus_{i=1}^{p} \mathbb{C}^{n_i \times n_i} \tag{2.7}$$

for later use, by mapping a matrix from the algebra into its block diagonal form and deleting the multiple blocks.

2.2 Examples of matrix $*$-algebras

This section is dedicated to the matrix $*$-algebras that we encounter later in this thesis.

2.2.1 Circulant matrices

A brief overview of circulant matrices follows. For more details the reader is referred to the review paper by Gray (2006).

A circulant matrix has the following form:

$$C == \begin{pmatrix} r_0 & r_1 & r_2 & . & . & r_{n-1} \\ r_{n-1} & r_0 & r_1 & r_2 & . & . \\ . & . & r_0 & r_1 & r_2 & . \\ . & . & . & . & . & . \\ r_2 & . & . & . & r_0 & r_1 \\ r_1 & r_2 & . & . & . & r_0 \end{pmatrix}, \tag{2.8}$$

that is,

$$C_{ij} := r_{(j-i) \bmod n}.$$

Sums, products, and conjugate transposes of such matrices remain circulant. As a consequence, the circulant matrices form a matrix ∗-algebra. Moreover, the multiplicative operation is commutative so this is a commutative matrix ∗-algebra. In addition, the eigenvalues of such matrices can easily be found exactly. Moreover, they share a common set of eigenvectors, given by the columns of the *discrete Fourier transform matrix*:

$$Q_{ij} := \frac{1}{\sqrt{n}} e^{\frac{-2\pi ij\sqrt{-1}}{n}}, \quad (i, j = 0, ..., n-1).$$

We have $Q^*Q = I$, and if C is a circulant matrix, then Q^*CQ is a diagonal matrix.

The eigenvalues of C are given by

$$\lambda_m(C) = r_0 + \sum_{k=1}^{n-1} r_k e^{-2\pi\sqrt{-1}mk/n}, \quad (m = 0, \ldots n-1).$$

The set of *symmetric* circulant matrices also forms a matrix ∗-algebra. In this case the closed-form expression for the eigenvalues, when n is odd, reduces to

$$\lambda_m(C) = r_0 + \sum_{k=1}^{(n-1)/2} 2r_k \cos(2\pi mk/n), \quad (m = 0, \ldots, n-1), \tag{2.9}$$

and when n is even we have

$$\lambda_m(C) = r_0 + \sum_{k=1}^{n/2-1} 2r_k \cos(2\pi mk/n) + r_{n/2}\cos(m\pi), \quad (m = 0, \ldots, n-1). \tag{2.10}$$

Furthermore, we may construct a basis $B_1, \ldots B_{\lfloor n/2 \rfloor}$ for the symmetric circulant matrices as follows. For each $i = 1, \ldots, \lfloor n/2 \rfloor$ define the entries of B_i by setting in (2.8) $r_i = r_{n-i} = 1$ and all the other r_j's to zero. We set $B_0 = I_n$. The positions of the nonzero entries in matrix B_i are sometimes called the *i-th stripe*, and we will use this terminology. Then, using (2.9) and (2.10), the eigenvalues of the basis matrices are

$$\lambda_m(B_k) = 2\cos(2\pi mk/n), \quad (m = 0, \ldots, n-1,\ k = 0, \ldots, \lfloor n/2 \rfloor). \tag{2.11}$$

Also note that, in view of Definition 2.1.2, the basis of the circulant matrices forms an association scheme.

2.2.2 Matrix $*$-algebras from permutation groups

Let $\mathcal{G} \subseteq \mathcal{S}_n$ be a subgroup of the symmetric group on n elements. With every element $\pi \in \mathcal{G}$ one can associate a permutation matrix $P_\pi \in \mathbb{C}^{n \times n}$ defined as follows:

$$(P_\pi)_{ij} = \begin{cases} 1 & \text{if } \pi(i) = j \\ 0 & \text{otherwise.} \end{cases} \quad (i, j = 1, \ldots, n). \tag{2.12}$$

Notice that

$$P_\pi^* = P_\pi^T = P_{\pi^{-1}}.$$

Moreover, for all $\pi, \rho \in \mathcal{G}$, if $\pi\rho$ denotes the permutation $\pi\rho(i) := \pi(\rho(i))$, we have

$$P_{\pi\rho} = P_\pi P_\rho \quad \text{and} \quad P_{\pi^{-1}} = P_\pi^{-1}.$$

This means that the mapping $\pi \mapsto P_\pi$ defines a *representation* of $\mathcal{G}$, called *the orthogonal representation.*

Definition 2.2.1. *Let $G = (V, E)$ be a graph. The automorphism group of G, denoted aut(G), is defined by those permutations of the vertices that preserve the adjacency structure of the graph.*

The orthogonal representation of aut(G) consists of the permutation matrices P having the property

$$P^T A P = A,$$

where A is the adjacency matrix of G.

Definition 2.2.2. *We define the automorphism group of a given $M \in \mathbb{C}^{n\times n}$ as*

$$aut(M) := \{\pi \in \mathcal{S}_n \mid M_{ij} = M_{\pi(i)\pi(j)}\ \forall i, j\}$$

In the same vein, given a matrix M, the orthogonal representation of aut(M) consists of the permutation matrices P having the property

$$P^T MP = M.$$

Throughout this thesis we will deal only with such orthogonal representations of permutation groups. Therefore, we will start with the set of permutation matrices defined in (2.12) and will refer to this set as the group $\mathcal{G}$. That is, in what follows we will not distinguish between the group and its orthogonal representation.

Definition 2.2.3. *The orbit of an element $i \in \{1, \dots n\}$ under the action of the group $\mathcal{G}$ is the set*

$$\{j \mid j = \pi(i) \quad for\ some\ \pi \in \mathcal{G}\}.$$

Similarly, the 2-orbit or orbital of an element $(i_1, i_2) \in \{1, \dots n\} \times \{1, \dots n\}$ under the action of the group $\mathcal{G}$ is the set

$$\{(j_1, j_2) \mid (j_1, j_2) = (\pi(i_1), \pi(i_2)) \quad for\ some\ \pi \in \mathcal{G}\}.$$

We can easily see that two distinct elements $i, j \in \{1, \dots n\}$ either have the same orbit or disjoint orbits under the action of $\mathcal{G}$.

Definition 2.2.4. *The centralizer ring (commutant) of the group $\mathcal{G}$ is defined as:*

$$\mathcal{A}_{\mathcal{G}} := \{Y \in \mathbb{R}^{n\times n} \mid Y = \frac{1}{|\mathcal{G}|}\sum_{P\in\mathcal{G}} P^T XP,\ \ X \in \mathbb{R}^{n\times n}\}. \tag{2.13}$$

The linear mapping $X \mapsto R(X) := \frac{1}{|\mathcal{G}|}\sum_{P\in\mathcal{G}} P^T XP,\ \ X \in \mathbb{R}^{n\times n}$ is called the *group average* or *Reynolds operator.*

Theorem 2.2.1. *We have the following equivalent formulation of $\mathcal{A}_{\mathcal{G}}$:*

$$\mathcal{A}_{\mathcal{G}} = \{Y \in \mathbb{R}^{n\times n} \mid PY = YP \quad \forall P \in \mathcal{G}\}. \tag{2.14}$$

Proof. Let us denote the set in (2.13) by $\mathcal{A}_1$ and the set in (2.14) by $\mathcal{A}_2$. We first show that $\mathcal{A}_1 \subseteq \mathcal{A}_2$.

Let $Y \in \mathcal{A}_1$. Then $Y = \frac{1}{|\mathcal{G}|}\sum_{P\in\mathcal{G}} P^T X P$ for some $X \in \mathbb{R}^{n\times n}$. We have to prove that $\overline{P}Y = Y\overline{P}$ $\forall \overline{P} \in \mathcal{G}$. The condition $\overline{P}Y = Y\overline{P}$ $\forall \overline{P} \in \mathcal{G}$ is equivalent to

$$\overline{P}\frac{1}{|\mathcal{G}|}\sum_{P\in\mathcal{G}} P^T X P = \left(\frac{1}{|\mathcal{G}|}\sum_{P\in\mathcal{G}} P^T X P\right)\overline{P} \quad \forall \overline{P} \in \mathcal{G},$$

and therefore to

$$\sum_{P\in\mathcal{G}} \overline{P}P^T X P = \sum_{P\in\mathcal{G}} P^T X P\overline{P} \quad \forall \overline{P} \in \mathcal{G}.$$

Now fix a permutation matrix $\overline{P} \in \mathcal{G}$. If we write $\alpha = |\mathcal{G}|$ and $\mathcal{G} = \{P_1, ..., P_\alpha\}$, then the equality above can be rewritten as

$$\overline{P}P_1^T X P_1 + ... + \overline{P}P_\alpha^T X P_\alpha = P_1^T X P_1\overline{P} + ... + P_\alpha^T X P_\alpha\overline{P}. \tag{2.15}$$

We want to show that these two sums have the same terms (not necessarily in the same order). Note that for any given $j \in \{1, ..., \alpha\}$ there exists $j^* \in \{1, ..., \alpha\}$ such that $\overline{P}P_j^T = P_{j^*}^T$. Then, to obtain the equality in (2.15) we still need to prove that $P_{j^*}\overline{P} = P_j$, but this is immediate since $P_{j^*}^T P_{j^*} = I = P_j^T P_j$.

Conversely, let $X \in \mathcal{A}_2$. This means $X \in \mathbb{R}^{n\times n}$ and $PX = XP$ $\forall P \in \mathcal{G}$. We have $X = P^T X P$ $\forall P \in \mathcal{G}$. Summing over all permutation matrices we obtain $X = \frac{1}{|\mathcal{G}|}\sum_{P\in\mathcal{G}} P^T X P$, hence $X \in \mathcal{A}_1$. $\square$

We will repeatedly use the following property of the Reynolds operator:

$$\operatorname{trace}(R(X)Y) = \operatorname{trace}(R(Y)X), \quad \forall X, Y \in \mathbb{R}^{n\times n}. \tag{2.16}$$

An important observation is that the Reynolds operator gives us the 2-orbits of elements of $\{1, \ldots n\} \times \{1, \ldots n\}$ under the action of $\mathcal{G}$. The 2-orbit of an element, say (i, j), corresponds to the nonzero entries of the matrix

$$\frac{1}{|\mathcal{G}|}\sum_{P\in\mathcal{G}} P^T e_i e_j^T P.$$

The 2-orbits define a coherent configuration, and the corresponding algebra is $\mathcal{A}_\mathcal{G}$.

Theorem 2.2.2. *$\mathcal{A}_\mathcal{G}$ is a matrix $*$-algebra. Moreover, there exists a basis $B_1, \ldots, B_d$ of $\mathcal{A}_\mathcal{G}$ with the following properties:*

(1) $\sum_{i \in \mathcal{I}} B_i = I$ *for some index set* $\mathcal{I} \subset \{1, \ldots, d\}$*;*

(2) B_i *is a 0-1 matrix for all* i*;*

(3) $\sum_{i=1}^{d} B_i = J$*;*

(4) Given i*,* $B_i = B_{i^*}^T$ *for some* i^**.*

Such a basis is given by the 2-orbits of $\mathcal{G}$*.*

Proof. We first show that $\mathcal{A}_{\mathcal{G}}$ is a matrix $*$-algebra. Let $A, B \in \mathcal{A}$ and $\alpha, \beta \in \mathbb{R}$, $P \in \mathcal{G}$. Then, from (2.14),

$$(\alpha A + \beta B)P = \alpha AP + \beta BP = \alpha PA + \beta PB = P(\alpha A + \beta B).$$

We conclude that $\mathcal{A}_{\mathcal{G}}$ is a linear subspace of $\mathbb{R}^{n \times n}$.

Further, since $P^{-1} = P^T$ we have $A = PAP^T$. Multiplying this member by the member with $BP = PB$ yields $ABP = PAP^TPB$. Thus, $ABP = PAB$ and we conclude that $AB \in \mathcal{A}$. Using $PP^T = I$ and $(AP)^* = P^*A^*$ we obtain $A^*P = PA^*$, which proves that the centralizer ring is also closed under conjugation.

We will now construct the required basis as the image under the Reynolds operator of the standard basis of $\mathbb{R}^{n \times n}$, i.e., from the matrices

$$\frac{1}{|\mathcal{G}|} \sum_{P \in \mathcal{G}} P^T e_i e_j^T P.$$

Since the orbits of $e_{i_1} e_{j1}^T$ and $e_{i_2} e_{j2}^T$ under the action of $\mathcal{G}$ are either the same or disjoint, we can define the 0-1 basis matrices by summing over the distinct (say d) 2-orbits:

$$B_k = \sum_{(i,j)\ has\ 2-orbit\ k} \frac{1}{|\mathcal{G}|} \sum_{P \in \mathcal{G}} P^T e_i e_j^T P \quad (k = 1, \ldots d).$$

Moreover, since $\sum_{i,j=1}^{n} e_i e_j^T = J$ we have

$$\sum_{i,j=1}^{n} \frac{1}{|\mathcal{G}|} \sum_{P \in \mathcal{G}} P^T e_i e_j^T P = J,$$

which means

$$\sum_{k=1}^{d} B_k = J.$$

Moreover,

$$\sum_{k\in\mathcal{I}} \sum_{(i,i)\ has\ 2-orbit\ k} \frac{1}{|\mathcal{G}|} \sum_{P\in\mathcal{G}} P^T e_i e_i^T P = I,$$

which proves the first property.

The last property follows from the observation that $(B_k)_{ij} = (B_k)_{ij}^T$, and this concludes the proof of the theorem. □

To conclude, starting from a permutation group $\mathcal{G} \subseteq \mathcal{S}_n$, we have constructed a matrix $*$-algebra $\mathcal{A}_{\mathcal{G}}$.

Let us consider the following example.

Example 2.2.1. *Consider the 5-cycle (pentagon), denoted C_5. The automorphism group of C_5 is the so-called* dihedral group *on 5 elements and has order $|aut(C_5)| = 10$. The orbits of $\{1,2,3,4,5\} \times \{1,2,3,4,5\}$ under the action of $\operatorname{aut}(C_5)$ correspond to the matrices:*

$$B_1 = \begin{pmatrix} 1&0&0&0&0\\ 0&1&0&0&0\\ 0&0&1&0&0\\ 0&0&0&1&0\\ 0&0&0&0&1 \end{pmatrix},\ B_2 = \begin{pmatrix} 0&1&0&0&1\\ 1&0&1&0&0\\ 0&1&0&1&0\\ 0&0&1&0&1\\ 1&0&0&1&0 \end{pmatrix},\ B_3 = \begin{pmatrix} 0&0&1&1&0\\ 0&0&0&1&1\\ 1&0&0&0&1\\ 1&1&0&0&0\\ 0&1&1&0&0 \end{pmatrix}.$$

Notice that B_1 has nonzero entries on the positions corresponding to the orbit of the pair $(1,1)$, B_2 to the orbit of the pair $(1,2)$, and B_3 to the orbit of the pair $(1,3)$. Thus, the centralizer ring of $aut(C_5)$ is a matrix $*$-algebra of dimension 3.

2.2.3 Coherent algebras containing given matrices

We consider now a given set of matrices $\{A_0, \ldots, A_m\}$. The goal is to identify the smallest coherent matrix $*$-algebra (say $\mathcal{A}$) that contains these matrices.

One approach is to assume that the multiplicative matrix group $\mathcal{G} := \bigcap_{i=0}^m \operatorname{aut}(A_i)$ is nontrivial. Then, as before, we may take $\mathcal{A}$ as the commutant of the permutation representation of $\mathcal{G}$. According to Theorem 2.2.2 the matrix $*$-algebra $\mathcal{A}$ will have a basis of 0-1 matrices (coherent configuration) given by the orbits of $\mathcal{G}$. However, this construction does not guarantee that we obtain the smallest coherent matrix $*$-algebra.

One way to obtain $\mathcal{A}$ is via the Weisfeiler-Leman algorithm; see Babel, Baumann, Lüdecke, and Tinhofer (1997). Algorithm 1 and Algorithm 2 briefly show how one can use the Weisfeiler-Leman approach to identify the smallest coherent matrix $*$-algebra containing $\{A_0, \ldots, A_m\}$.

Algorithm 1 (Weisfeiler-Leman)

INPUT: A matrix M, whose identical entries define a partition $B_1, ..., B_d$ that satisfies properties (1)–(4) in Theorem 2.2.2.
(i) Replace the entries of the matrix M by noncommuting variable symbols $s_0, \ldots, s_{k_1}$, and call it M_{k_1}.
(ii) Compute $M_{k_1}^2$ and denote its distinct entries by $s_0, ..., s_{k_2}$.
(iii) Repeat the computation until no new symbols appear; and let $s_0, ..., s_{k_n}$ denote the symbols in the configuration of M_{k_n}.
(iv) Identify a coherent configuration by constructing $k_n + 1$ 0-1 matrices having 1 in the position of the common symbols and zero elsewhere.
OUTPUT: The coherent configuration of the smallest coherent algebra that contains M.

Algorithm 2 Smallest coherent algebra containing given matrices

INPUT: The set of matrices $\{A_0, \ldots, A_m\}$.
(i) Take a random combination of the input matrices, call it M.
(ii) Perform the Weisfeiler-Leman algorithm on matrix M.
OUTPUT: A coherent configuration of the smallest coherent algebra that contains $\{A_0, \ldots, A_m\}$.

Example 2.2.2. *We illustrate Algorithm 1 for the set of 5×5 circulant matrices. First, notice that any random linear combination of circulant matrices remains circulant. To avoid triviality (completion after one iteration) we will consider two equal entries in the first row of the matrix from (2.8). Hence, consider the following matrix:*

$$M = \begin{pmatrix} 0 & 1 & 2 & 1 & 3 \\ 3 & 0 & 1 & 2 & 1 \\ 1 & 3 & 0 & 1 & 2 \\ 2 & 1 & 3 & 0 & 1 \\ 1 & 2 & 1 & 3 & 0 \end{pmatrix}$$

Replace the entries of the matrix M by noncommuting variable symbols and obtain:

$$M_{k_1} = \begin{pmatrix} s_0 & s_1 & s_2 & s_1 & s_3 \\ s_3 & s_0 & s_1 & s_2 & s_1 \\ s_1 & s_3 & s_0 & s_1 & s_2 \\ s_2 & s_1 & s_3 & s_0 & s_1 \\ s_1 & s_2 & s_1 & s_3 & s_0 \end{pmatrix}.$$

Further,

$$M_{k_2} := M_{k_1}^2 = \begin{pmatrix} s_0 & s_1 & s_2 & s_3 & s_4 \\ s_4 & s_0 & s_1 & s_2 & s_3 \\ s_3 & s_4 & s_0 & s_1 & s_2 \\ s_2 & s_3 & s_4 & s_0 & s_1 \\ s_1 & s_2 & s_3 & s_4 & s_0 \end{pmatrix},$$

where the new symbols s_0, s_1, s_2, s_3, s_4 *are given by*

$$\begin{aligned} s_0 &\leftarrow s_0^2 + s_1 s_3 + s_2 s_1 + s_1 s_2 + s_3 s_1, \\ s_1 &\leftarrow s_0 s_1 + s_1 s_0 + s_2 s_3 + s_1^2 + s_3 s_2, \\ s_2 &\leftarrow s_0 s_2 + s_1^2 + s_2 s_0 + s_1 s_3 + s_3 s_1, \\ s_3 &\leftarrow s_0 s_1 + s_1 s_2 + s_2 s_1 + s_1 s_0 + s_3^2, \\ s_4 &\leftarrow s_0 s_3 + 2 s_1^2 + s_2^2 + s_3 s_0. \end{aligned}$$

Notice that we obtained one extra symbol, s_4*, so we have to compute* $M_{k_3} := M_{k_2}^2$*. This computation is carried out in a similar way and the resulting* M_{k_3} *does not introduce any new symbols so the algorithm stops.*

We identify the coherent configuration of dimension 5 (i.e., $\{B_0, ..., B_4\}$*) from the formulation* $M_{k_3} = M_{k_2} = \sum_{i=0}^{4} s_i B_i$*.*

We have

$$B_0 = \begin{pmatrix} 1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1 \end{pmatrix}, B_1 = \begin{pmatrix} 0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\\1&0&0&0&0 \end{pmatrix}, B_2 = \begin{pmatrix} 0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\\1&0&0&0&0\\0&1&0&0&0 \end{pmatrix},$$

$$B_3 = \begin{pmatrix} 0&0&0&1&0\\0&0&0&0&1\\1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0 \end{pmatrix}, B_4 = \begin{pmatrix} 0&0&0&0&1\\1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0 \end{pmatrix}.$$

2.2.4 Coherent algebras associated with special graphs

We now give three more examples of matrix $*$-algebras that we will need later in this thesis (i.e., Chapter 5). Since the cardinality of the basis gives the dimension of the matrix $*$-algebra we will use the notion of *dimension* instead of *rank* when referring to the cardinality of a coherent configuration.

Example 2.2.3. *Consider the coherent configuration arising from $\mathcal{G} = aut(K_{m-1,m})$ associated with the complete bipartite graph $K_{m-1,m}$. The coherent configuration has dimension 6 and consists of the following matrices:*

$$A_1 = \begin{pmatrix} I_{m-1} & 0_{m-1\times m} \\ 0_{m\times m-1} & 0_{m\times m} \end{pmatrix},\ A_2 = \begin{pmatrix} J_{m-1} - I_{m-1} & 0_{m-1\times m} \\ 0_{m\times m-1} & 0_{m\times m} \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 0_{m-1\times m-1} & J_{m-1\times m} \\ 0_{m\times m-1} & 0_{m\times m} \end{pmatrix},\ A_4 = \begin{pmatrix} 0_{m-1\times m-1} & 0_{m-1\times m} \\ J_{m\times m-1} & 0_{m\times m} \end{pmatrix},$$

$$A_5 = \begin{pmatrix} 0_{m-1\times m-1} & 0_{m-1\times m} \\ 0_{m\times m-1} & I_{m\times m} \end{pmatrix},\ A_6 = \begin{pmatrix} 0_{m-1\times m-1} & 0_{m-1\times m} \\ 0_{m\times m-1} & J_m - I_m. \end{pmatrix},$$

and its complex span is isomorphic (as a $*$-algebra) to $\mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C}^{2\times 2}$. The relevant $*$-isomorphism, say ϕ (see (2.7)), satisfies:

$$\phi(A_1) = \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 1 & 0 \\ & & 0 & 0 \end{pmatrix},\ \phi(A_2) = \begin{pmatrix} -1 & & & \\ & 0 & & \\ & & m-2 & 0 \\ & & 0 & 0 \end{pmatrix},\ \phi(A_3) = \sqrt{(m-1)m}\begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & 1 \\ & & 0 & 0 \end{pmatrix},$$

$$\phi(A_4) = \sqrt{(m-1)m}\begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & 0 \\ & & 1 & 0 \end{pmatrix},\ \phi(A_5) = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 0 & 0 \\ & & 0 & 1 \end{pmatrix},\ \phi(A_6) = \begin{pmatrix} 0 & & & \\ & -1 & & \\ & & 0 & 0 \\ & & 0 & m-1 \end{pmatrix}.$$

Example 2.2.4. *Consider the following coherent configuration arising from $\mathcal{G} = aut(K_{m-1,\ldots,m})$ associated with the complete multipartite graph $K_{m-1,m,\ldots,m}$ (i.e., k-partition of cardinality given by indices) where each matrix contains k^2 blocks (block dimensions are given only for the first matrix; they can be deduced from the context):*

$$A_1 = \begin{pmatrix} I_{m-1} & 0_{m-1\times m} & 0_{m-1\times m} & \ldots & 0_{m-1\times m} \\ 0_{m\times m-1} & 0_{m\times m} & 0_{m\times m} & \ldots & 0_{m\times m} \\ 0_{m\times m-1} & 0_{m\times m} & 0_{m\times m} & \ldots & 0_{m\times m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_{m\times m-1} & 0_{m\times m} & 0_{m\times m} & \ldots & 0_{m\times m} \end{pmatrix},$$

$$A_2 = \begin{pmatrix} J-I & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix},\ A_3 = \begin{pmatrix} 0 & J & J & \ldots & J \\ 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix},$$

$$A_4 = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 \\ J & 0 & 0 & \ldots & 0 \\ J & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ J & 0 & 0 & \ldots & 0 \end{pmatrix},\ A_5 = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 \\ 0 & I & 0 & \ldots & 0 \\ 0 & 0 & I & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & I \end{pmatrix},$$

$$A_6 = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & J-I & 0 & \dots & 0 \\ 0 & 0 & J-I & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & J-I \end{pmatrix}, \; A_7 = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & J & \dots & J \\ 0 & J & 0 & \dots & J \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & J & J & \dots & 0 \end{pmatrix}.$$

The complex span is isomorphic to $\mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C}^{2\times 2}$*. The relevant* $*$*-isomorphism, say* ϕ*, satisfies:*

$$\phi(A_1) = \begin{pmatrix} 1 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & 1 & 0 \\ & & & 0 & 0 \end{pmatrix}, \; \phi(A_2) = \begin{pmatrix} -1 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & m-2 & 0 \\ & & & 0 & 0 \end{pmatrix},$$

$$\phi(A_3) = \sqrt{(k-1)m(m-1)} \begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & 0 & 1 \\ & & & 0 & 0 \end{pmatrix},$$

$$\phi(A_4) = \sqrt{(k-1)m(m-1)} \begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & 0 & 0 \\ & & & 1 & 0 \end{pmatrix},$$

$$\phi(A_5) = \begin{pmatrix} 0 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 0 & 0 \\ & & & 0 & 1 \end{pmatrix}, \; \phi(A_6) = \begin{pmatrix} 0 & & & & \\ & -1 & & & \\ & & m-1 & & \\ & & & 0 & 0 \\ & & & 0 & m-1 \end{pmatrix},$$

$$\phi(A_7) = m \begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & -1 & & \\ & & & 0 & 0 \\ & & & 0 & k-2 \end{pmatrix}.$$

Example 2.2.5. *Consider the commutative coherent configuration (i.e., association scheme) arising from* $\mathcal{G} = aut(K_{m,\dots,m})$ *associated with the complete multipartite graph* $K_{m,\dots,m}$*. The coherent configuration has dimension 3 and consists of the following matrices:*

$$A_1 = \begin{pmatrix} I_m & 0_m & \dots & 0_m \\ 0_m & I_m & \dots & 0_m \\ \vdots & \vdots & \ddots & \vdots \\ 0_m & 0_m & \dots & I_m \end{pmatrix}, \; A_2 = \begin{pmatrix} 0_m & J_m & \dots & J_m \\ J_m & 0_m & \dots & J_m \\ \vdots & \vdots & \ddots & \vdots \\ J_m & J_m & \dots & 0_m \end{pmatrix},$$

$$A_3 = \begin{pmatrix} J_m - I_m & 0_m & \dots & 0_m \\ 0_m & J_m - I_m & \dots & 0_m \\ \vdots & \vdots & \ddots & \vdots \\ 0_m & 0_m & \dots & J_m - I_m \end{pmatrix}.$$

Its complex span is isomorphic (as a ∗-algebra) to $\mathbb{C} \oplus \mathbb{C} \oplus \mathbb{C}$. The relevant ∗-isomorphism, say ϕ, satisfies:

$$\phi(A_1) = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix},\ \phi(A_2) = \begin{pmatrix} (k-1)m & & \\ & 0 & \\ & & -m \end{pmatrix},\ \phi(A_3) = \begin{pmatrix} m-1 & & \\ & -1 & \\ & & m-1 \end{pmatrix}.$$

A few remarks on the ∗-isomorphisms from the previous examples:

- The block structure of $\phi(A_i)$ can be deduced using (2.7), applying for example the algorithm in De Klerk, Dobre, and Pasechnik (2010) to block diagonalize the matrices A_i.
- The eigenvalues of A_i and $\phi(A_i)$ are the same up to multiplicities.
- One can verify that ϕ is a ∗-isomorphism by noticing that the multiplication tables of the $\phi(A_i)$'s and the A_i's are the same; and further by verifying $\phi(A_iA_j) = \phi(A_i)\phi(A_j)$, for any i, j, and $\phi(A_i^*) = \phi(A_i)^*$ for any i.

2.3 The regular *-representation of matrix ∗-algebras

In Chapter 3 we will not compute the Wedderburn decomposition of a given matrix $\mathbb{C}*$-algebra $\mathcal{A}$ directly. We will instead compute the Wedderburn decomposition of a faithful (i.e., isomorphic) representation of it, called the regular *-representation of $\mathcal{A}$. Of course, the end result is the same, since the Wedderburn decomposition is canonical, but this approach allows numerical computation with smaller matrices.

2.3.1 General facts

Definition 2.3.1 (see e.g., Section 1 in Etingof, Golberg, Hensel, Liu, Schwendner, Udovina, and Vaintrob (2009))**.** *A representation of an algebra $\mathcal{A}$ is a vector space V together with a homomorphism of algebras $\varphi : \mathcal{A} \mapsto \mathrm{End}(V)$, where $\mathrm{End}(V)$ denotes the set of endomorphisms from V to V.*

Definition 2.3.2. *When $V = \mathcal{A}$ and $\varphi : \mathcal{A} \mapsto \mathrm{End}(\mathcal{A})$ is given by $\varphi(A)Y = AY$ $\forall Y \in \mathcal{A}$, one obtains the regular representation of $\mathcal{A}$. Moreover, when $\mathcal{A}$ has an involution operation, say ∗, and $\varphi(A^*) = \varphi(A)^*$ $\forall\ A \in \mathcal{A}$, one obtains the regular *-representation.*

Note that in the definition above A^* is the involution of A, and $\varphi(A)^*$ is the adjoint of the linear operator $\varphi(A)$. We will use the notation $\varphi_A := \varphi(A)$, so $\varphi_A(Y) = AY$ $\forall Y \in \mathcal{A}$.

Assume now that $\mathcal{A}$ has an orthogonal basis of real matrices $B_1, \ldots, B_d \in \mathbb{R}^{n\times n}$, with $B_i^* \in \{B_1, ..., B_d\}$ for any $i = 1, ..., d$. This situation is not general, but it is usual for the applications in semidefinite programming that we will consider.

We normalize this basis with respect to the *Frobenius norm*:

$$D_i := \frac{1}{\sqrt{\operatorname{trace}(B_i^T B_i)}} B_i \quad (i = 1, \ldots, d),$$

and define multiplication parameters $\gamma_{i,j}^k$ via:

$$D_i D_j = \sum_{k=1}^{d} \gamma_{i,j}^k D_k, \tag{2.17}$$

and subsequently define the $d \times d$ matrices L_k $(k = 1, \ldots, d)$ via

$$(L_k)_{ij} = \gamma_{k,j}^i, \quad (i, j = 1, \ldots, d). \tag{2.18}$$

Lemma 2.3.1. *For any $k = 1, ..., d$, L_k is the matrix representation of the linear operator φ_{D_k} with respect to the basis $\{D_1, ..., D_d\}$.*

Proof. Since $D_k \in \mathcal{A}$, for any $k = 1, ..., d$ we have

$$\varphi_{D_k}(D_j) = D_k D_j = \sum_{i=1}^{d} (L_k)_{ij} D_i, \ \ (j = 1, ..., d),$$

which completes the proof. □

Therefore, we will work with the matrix representation of the linear operator φ_{D_k}. The matrices L_k form the basis of a matrix $*$-algebra, say $\mathcal{A}^{reg}$. We will abuse terminology slightly by calling $\mathcal{A}^{reg}$ the *regular *-representation of* $\mathcal{A}$ (with respect to the basis $\{D_1, ..., D_d\}$).

Theorem 2.3.2. *The bijective linear mapping $\Phi : \mathcal{A} \mapsto \mathcal{A}^{reg}$ such that $\Phi(D_k) = L_k$ $(k = 1, \ldots, d)$ defines a $*$-isomorphism from $\mathcal{A}$ to $\mathcal{A}^{reg}$. Thus, Φ is an algebra isomorphism with the additional property*

$$\Phi(A^*) = \Phi(A)^* \quad \forall A \in \mathcal{A}.$$

Proof. For any $Y \in \mathcal{A}$ we define as before the linear operator $\varphi_Y : \mathcal{A} \mapsto \mathcal{A}$ by

$$\varphi_Y(X) = YX \ \ \forall X \in \mathcal{A}. \tag{2.19}$$

Using Lemma 2.3.1 we have that $L_k := \Phi(D_k)$ is the matrix corresponding to the linear operator φ_{D_k} in the basis $D_1, ..., D_d$. Thus, for any $Y = \sum_k y_k D_k \in \mathcal{A}$, $\Phi(Y)$ is the matrix corresponding to the linear operator

$$\varphi_Y = \varphi_{\sum_k y_k D_k} = \sum_k y_k \varphi_{D_k}$$

in the basis $D_1, ..., D_d$.

Using (2.19) we have for any $Y, Z \in \mathcal{A}$:

$$\varphi_{YZ}(X) = YZX = Y(ZX) = \varphi_Y(\varphi_Z(X)) = (\varphi_Y \circ \varphi_Z)(X) \quad \forall X \in \mathcal{A}.$$

Therefore, for any $Y, Z \in \mathcal{A}$ we have $\Phi(YZ) = \Phi(Y)\Phi(Z)$. Thus, Φ is an algebra homomorphism.

$\Phi(Y) = 0$ implies that $YX = 0 \ \forall X \in \mathcal{A}$, and in particular we obtain $YY^* = 0$, which implies that $Y = 0$. Therefore, Φ is injective and by construction we conclude that it is a bijection.

We still need to show that Φ is a *-isomorphism (i.e., it preserves symmetry). To do so, we need to show that $\Phi(Y^*) = \Phi(Y)^*$.

On the one hand, by definition of φ_Y we have $\varphi_Y(D_j) = YD_j$; on the other hand, using the fact that $\Phi(Y)$ is the matrix of operator φ_Y in the basis $D_1, ..., D_d$ we obtain:

$$YD_j = \sum_{t=1}^{d} \Phi(Y)_{tj} D_t.$$

Using the orthonormality of the basis $D_1, ..., D_d$, in the above relation, we take the inner product with the matrices D_i and use the linearity of the operator. Hence,

$$\mathrm{trace}(D_i^T Y D_j) = \sum_{t=1}^{d} \Phi(Y)_{tj} \mathrm{trace}(D_i^T D_t) = \Phi(Y)_{ij}.$$

In the same way:

$$Y^* D_i = \sum_{t=1}^{d} \Phi(Y^*)_{ti} D_t$$

and we take the inner product with the matrices D_j. Notice that if $A \in \mathbb{C}^{n\times n}$ then $\text{trace}(A^*) = \overline{\text{trace}(A)}$. From the orthonormality of the basis, the right-hand side becomes $\Phi(Y^*)_{ji}$. Hence,

$$\Phi(Y^*)_{ji} = \text{trace}(D_j^T Y^* D_i) = \overline{\text{trace}(D_j^T Y^* D_i)^*} = \overline{\text{trace}(D_i^T Y D_j)} = \overline{\Phi(Y)_{ij}},$$

therefore the preservation of the symmetry is proved. □

Since Φ is a homomorphism, A and $\Phi(A)$ have the same eigenvalues (up to multiplicities) for all $A \in \mathcal{A}$. As a consequence, we have the following theorem.

Theorem 2.3.3. *Let $\{D_1, ..., D_d\}$ be an orthonormal basis of a matrix $*$-algebra $\mathcal{A}$, $\{L_1, ..., L_d\}$ the basis of the regular *-representation of $\mathcal{A}$ (i.e., $\mathcal{A}^{reg}$) as defined in (2.18), and $x \in \mathbb{R}^d$. We have*

$$\sum_{i=1}^{d} x_i D_i \succeq 0 \iff \sum_{i=1}^{d} x_i L_i \succeq 0.$$

Example 2.3.1. *We revisit Example 2.2.1. Recall that $n = 5$ and $d = 3$ for this example. First, we normalize the basis B_1, B_2, B_3 to get*

$$D_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1&0&0&0&0\\ 0&1&0&0&0\\ 0&0&1&0&0\\ 0&0&0&1&0\\ 0&0&0&0&1 \end{pmatrix},\ D_2 = \frac{1}{\sqrt{10}}\begin{pmatrix} 0&1&0&0&1\\ 1&0&1&0&0\\ 0&1&0&1&0\\ 0&0&1&0&1\\ 1&0&0&1&0 \end{pmatrix},\ D_3 = \frac{1}{\sqrt{10}}\begin{pmatrix} 0&0&1&1&0\\ 0&0&0&1&1\\ 1&0&0&0&1\\ 1&1&0&0&0\\ 0&1&1&0&0 \end{pmatrix}.$$

Then, Table 2.1 will give us the coefficients $\gamma_{i,j}^k$, $i, j, k = 1, \ldots 3$.

	D_1	D_2	D_3
D_1	$\frac{1}{\sqrt{5}}D_1$	$\frac{1}{\sqrt{5}}D_2$	$\frac{1}{\sqrt{5}}D_3$
D_2	$\frac{1}{\sqrt{5}}D_2$	$\frac{1}{\sqrt{5}}D_1 + \frac{1}{\sqrt{10}}D_3$	$\frac{1}{\sqrt{10}}(D_2 + D_3)$
D_3	$\frac{1}{\sqrt{5}}D_3$	$\frac{1}{\sqrt{10}}(D_2 + D_3)$	$\frac{1}{\sqrt{5}}D_1 + \frac{1}{\sqrt{10}}D_2$

Table 2.1: Multiplication table of normalized matrices from Example 2.2.1.

Further, using (2.17) and (2.18), we can easily compute by hand the matrices L_1, L_2, L_3, that form the basis of the regular $*$-representation:

$$L_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1&0&0\\ 0&1&0\\ 0&0&1 \end{pmatrix},\ L_2 = \begin{pmatrix} 0&\frac{1}{\sqrt{5}}&0\\ \frac{1}{\sqrt{5}}&0&\frac{1}{\sqrt{10}}\\ 0&\frac{1}{\sqrt{10}}&\frac{1}{\sqrt{10}} \end{pmatrix},\ L_3 = \begin{pmatrix} 0&0&\frac{1}{\sqrt{5}}\\ 0&\frac{1}{\sqrt{10}}&\frac{1}{\sqrt{10}}\\ \frac{1}{\sqrt{5}}&\frac{1}{\sqrt{10}}&0 \end{pmatrix}.$$

Notice that in this toy example we have reduced the size of the basis matrices from $n = 5$ to $d = 3$.

2.3.2 A change of basis

By Wedderburn's theorem, any matrix $\mathbb{C}*$-algebra $\mathcal{A}$ that contains the identity takes the form

$$Q^*\mathcal{A}Q = \oplus_{i=1}^{t} t_i \odot \mathbb{C}^{n_i\times n_i}, \tag{2.20}$$

for some integers t, t_i, and n_i $(i = 1,\dots,t)$, and some unitary Q.

Our further goal is to show that the Wedderburn decomposition of $\mathcal{A}^{reg}$ has a special structure that does not depend on the values t_i $(i = 1,\dots,t)$. To this end, the lemmas in this subsection show how the regular *-representation behaves when the orthonormal basis of the matrix $*$-algebra $\mathcal{A}$ is changed.

Lemma 2.3.4. *The regular *-representations of $\mathcal{A}$ and $Q^*\mathcal{A}Q$ are the same.*

Proof. Denote by $\mathcal{A}_Q$ the algebra after block diagonalization.

We have that $\{Q^*D_1Q, ..., Q^*D_dQ\}$ is a basis for $\mathcal{A}_Q$. We will prove that applying the regular *-representation to both $\mathcal{A}$ and $\mathcal{A}_Q$ yields the same matrices denoted earlier in this section by $L_1, ..., L_d$.

If we denote $D_i^{'} := Q^*D_iQ$, then from (2.17), by multiplying with Q^* and Q to the left and right respectively we obtain:

$$Q^*D_iD_jQ = \sum_{k=1}^{d} \gamma_{i,j}^{k} Q^*D_kQ.$$

Further, since Q is unitary, we have

$$Q^*D_iQQ^*D_jQ = \sum_{k=1}^{d} \gamma_{i,j}^{k} Q^*D_kQ,$$

and using the earlier notation

$$D_i^{'}D_j^{'} = \sum_{k=1}^{d} \gamma_{i,j}^{k} D_k^{'},$$

which proves that we have the same values $\gamma_{i,j}^{k}$ so we obtain the same regular *-representation for both $\mathcal{A}$ and $\mathcal{A}_Q$. $\square$

This implies that, when studying $\mathcal{A}^{reg}$, we may assume without loss of generality that $\mathcal{A}$ takes the form

$$\mathcal{A} = \oplus_{i=1}^{t} t_i \odot \mathbb{C}^{n_i\times n_i}.$$

Lemma 2.3.5 (see e.g., Sections 0.1.0 and 1.0.1 in Horn and Johnson (1990))**.** *Let V be a vector space of dimension d:=dim(V), let $L : V \mapsto V$ be a linear operator, and $\mathcal{B} = \{B_1, ..., B_d\}$, $\mathcal{B}' := \{B_1', ..., B_d'\}$ two bases of V. Then there exists a matrix $S \in \mathbb{R}^{d \times d}$, independent of L, such that*

$$M_{\mathcal{B}}^L = S^{-1} M_{\mathcal{B}'}^L S,$$

where $M_{\mathcal{B}}^L$ is the matrix representation of L with respect to basis $\mathcal{B}$ and S is the transition matrix from $\mathcal{B}$ to $\mathcal{B}'$.

Corollary 2.3.6. *Let V be a vector space of dimension d:=dim(V), let $L : V \mapsto V$ be a linear operator, and $\mathcal{B} = \{D_1, ..., D_d\}$, $\mathcal{B}' := \{D_1', ..., D_d'\}$ two orthonormal bases of V. Then there exists a unitary matrix $Q \in \mathbb{R}^{d \times d}$, independent of L, such that*

$$M_{\mathcal{B}}^L = Q^* M_{\mathcal{B}'}^L Q,$$

where $M_{\mathcal{B}}^L$ is the matrix representation of L with respect to basis $\mathcal{B}$, and Q is the unitary transition matrix from $\mathcal{B}$ to $\mathcal{B}'$.

Proof. Let B denote the unitary matrix containing the orthonormal vectors of $\mathcal{B}$, and B' denote the unitary matrix containing the orthonormal vectors of $\mathcal{B}'$. If Q is the transition matrix, then $B = Q^* B'$. Since both B and B' are unitary matrices, it follows that Q is also unitary. Using Lemma 2.3.5 we conclude the proof. □

Lemma 2.3.7. *Let $\mathcal{A}^{reg}$ be the regular *-representation of $\mathcal{A}$ with respect to the orthonormal basis $\{D_1, ..., D_d\}$, and let $\mathcal{A}'^{\,reg}$ be the regular *-representation of $\mathcal{A}$ with respect to the orthonormal basis $\{D_1', ..., D_d'\}$. Then there exists a unitary matrix Q such that*

$$\mathcal{A}^{reg} = Q^* \mathcal{A}'^{\,reg} Q.$$

Proof. Define as before the linear mappings $\Phi : \mathcal{A} \mapsto \mathcal{A}^{reg}$ such that $\Phi(D_k) = L_k$ $(k = 1, \ldots, d)$, and $\Phi' : \mathcal{A} \mapsto \mathcal{A}'^{\,reg}$ such that $\Phi'(D_k') = L_k'$ $(k = 1, \ldots, d)$. Then we have

$$\begin{aligned} \mathcal{A}^{reg} &= \left\{ \sum_{k=1}^{d} \alpha_k L_k \mid \alpha_k \in \mathbb{C} \right\} \text{ and} \\ \mathcal{A}'^{\,reg} &= \left\{ \sum_{k=1}^{d} \alpha_k' L_k' \mid \alpha_k' \in \mathbb{C} \right\}, \end{aligned}$$

where, by Lemma 2.3.1, $L_k = \Phi(D_k)$ is the matrix corresponding to the linear operator φ_{D_k} in the basis $\{D_1, ..., D_d\}$, and $L_k^{'} = \Phi^{'}(D_k^{'})$ is the matrix corresponding to the linear operator $\varphi_{D_k^{'}}$ in the basis $\{D_1^{'}, ..., D_d^{'}\}$.

Thus, for any $A = \sum_k \alpha_k D_k \in \mathcal{A}$, $\Phi(A)$ is the matrix corresponding to the linear operator

$$\varphi_A = \varphi_{\sum_k \alpha_k D_k} = \sum_k \alpha_k \varphi_{D_k}$$

in the basis $D_1, ..., D_d$. Moreover, if we write $A = \sum_k \alpha_k^{'} D_k^{'} \in \mathcal{A}$, then $\Phi^{'}(A)$ is the matrix corresponding to the linear operator

$$\varphi_A = \varphi_{\sum_k \alpha_k^{'} D_k^{'}} = \sum_k \alpha_k^{'} \varphi_{D_k^{'}}$$

in the basis $D_1^{'}, ..., D_d^{'}$.

By Corollary 2.3.6, if $A \in \mathcal{A}$, the matrix representations of φ_A with respect to the two orthonormal bases $\{D_1, ..., D_d\}$ and $\{D_1^{'}, ..., D_d^{'}\}$ are related via

$$\Phi(A) = Q^* \Phi^{'}(A) Q,$$

where Q is some orthonormal matrix that does not depend on A. This concludes the proof. $\square$

2.3.3 Wedderburn decomposition of regular *-representation

Lemma 2.3.8. *Let t and n be given integers. The regular *-representation of $t \odot \mathbb{C}^{n\times n}$ is equivalent to $n \odot \mathbb{C}^{n\times n}$, for the standard basis.*

Proof. The standard basis of $t \odot \mathbb{C}^{n\times n}$ clearly has n^2 elements since we have t repeated blocks. Let

$$D_{i_1 i_2} := \frac{1}{\sqrt{t}} I_t \otimes e_{i_1} e_{i_2}^T, \quad (i_1, i_2 = 1, \ldots, n)$$

denote the normalized basis matrices. Its regular *-representation will consist of n^2 dimensional matrices, say $L_{i_1 i_2}$, $(i_1, i_2 = 1, \ldots, n)$.

We will show that for all i_1, i_2 we have $L_{i_1 i_2} = \frac{1}{\sqrt{t}} P^T (I_n \otimes (e_{i_2} e_{i_1}^T)) P$, for some permutation matrix P, and the lemma will therefore be proved.

To this end, for $i_1, i_2 \in \{1, \ldots, n\}$ let us define $E^{(i_1 i_2)} := e_{i_1} e_{i_2}^T$. Then, using (2.17), we have

$$\frac{1}{t}(I_t \otimes E^{(i_1 i_2)})(I_t \otimes E^{(j_1 j_2)}) = \sum_{k_1, k_2 = 1}^{n} \gamma_{(i_1 i_2),(j_1 j_2)}^{(k_1 k_2)} \frac{1}{\sqrt{t}} I_t \otimes E^{(k_1 k_2)}, \quad (i_1, i_2, j_1, j_2 = 1, \ldots, n),$$

for some scalars $\gamma^{(k_1k_2)}_{(i_1i_2),(j_1j_2)}$. This is equivalent to

$$\frac{1}{\sqrt{t}} I_t \otimes (E^{(i_1i_2)}E^{(j_1j_2)}) = I_t \otimes \sum_{k_1,k_2=1}^{n} \gamma^{(k_1k_2)}_{(i_1i_2),(j_1j_2)} E^{(k_1k_2)}, \quad (i_1, i_2, j_1, j_2 = 1, \ldots, n).$$

This yields:

$$\frac{1}{\sqrt{t}} E^{(i_1i_2)}E^{(j_1j_2)} = \sum_{k_1,k_2=1}^{n} \gamma^{(k_1k_2)}_{(i_1i_2),(j_1j_2)} E^{(k_1k_2)}, \quad (i_1, i_2, j_1, j_2 = 1, \ldots, n).$$

Since

$$E^{(i_1i_2)}E^{(j_1j_2)} = e_{i_1}e_{i_2}^T e_{j_1}e_{j_2}^T = \delta_{i_2j_1}E^{(i_1j_2)},$$

we have

$$\begin{aligned} \gamma^{(k_1k_2)}_{(i_1i_2),(j_1j_2)} &= \frac{1}{\sqrt{t}}\delta_{i_2j_1}\delta_{i_1k_1}\delta_{j_2k_2} \\ &= \begin{cases} \frac{1}{\sqrt{t}} & \text{if } k_1 = i_1,\ i_2 = j_1\ k_2 = j_2 \\ 0 & \text{else.} \end{cases} \end{aligned}$$

Using

$$(L_{i_1i_2})_{(j_1j_2),(k_1k_2)} = \gamma^{(i_1i_2)}_{(k_1k_2),(j_1j_2)}$$

we obtain

$$L_{i_1i_2} = \frac{1}{\sqrt{t}}(e_{i_2}e_{i_1}^T) \otimes I_n.$$

Following (2.3) we obtain, for a suitable permutation matrix P, $L_{i_1i_2} = \frac{1}{\sqrt{t}}P^T(I_n \otimes (e_{i_2}e_{i_1}^T))P$, and this concludes the proof. □

Lemma 2.3.9. *Let t and n be given integers. The regular *-representation of $t \odot \mathbb{C}^{n\times n}$ is equivalent to $n \odot \mathbb{C}^{n\times n}$, for any choice of orthonormal basis.*

Proof. By Lemma 2.3.8 the regular *-representation of $t \odot \mathbb{C}^{n\times n}$ is equivalent to $n \odot \mathbb{C}^{n\times n}$ when using the standard basis $I_t \otimes (e_{i_1}e_{i_2}^T)$. Lemma 2.3.7 completes the proof. □

Lemma 2.3.10. *Let $\mathcal{A}_\alpha$ be matrix $*$-algebras and let $\mathcal{A}_\alpha^{reg}$ denote their regular *-representations, for $\alpha = 1, ..., t$. The regular *-representation of $\oplus_{\alpha=1}^t \mathcal{A}_\alpha$ is equivalent to $\oplus_{\alpha=1}^t \mathcal{A}_\alpha^{reg}$.*

Proof. Let $\{D_1^\alpha, \ldots, D_{d_\alpha}^\alpha\}$ denote a given orthonormal basis of $\mathcal{A}_\alpha$ for each $\alpha = 1, \ldots, t$. Denote the regular *-representation of each $\mathcal{A}_\alpha$ by $\mathcal{A}_\alpha^{reg}$, with basis $\{L_1^\alpha, \ldots, L_{d_\alpha}^\alpha\}$. Let $d := \sum_{\alpha=1}^t d_\alpha$ be the dimension of $\mathcal{A} := \oplus_{\alpha=1}^t \mathcal{A}_\alpha$. We now construct an orthonormal basis, say $\{D_1, \ldots, D_d\}$ of $\mathcal{A}$. Each matrix D_i will be block diagonal with exactly one nonzero block given by D_j^α for some $\alpha \in \{1, \ldots, t\}$ and $j \in \{1, \ldots, d_\alpha\}$. The position of this nonzero block will correspond to the "position"' of $\mathcal{A}_\alpha$ in the direct sum $\oplus_{\alpha=1}^t \mathcal{A}_\alpha$.

The exact construction is as follows: matrices $D_1, \ldots, D_{d_1}$ are formed from $D_1^1, \ldots, D_{d_1}^1$ respectively, matrices $D_{d_1+1}, \ldots, D_{d_1+d_2}$ are formed from $D_1^2, \ldots, D_{d_2}^2$ respectively, etc. If we denote the regular $*$-representation of $\mathcal{A}$ by $\mathcal{A}^{reg}$, with basis $\{L_1, \ldots, L_d\} \subset \mathbb{C}^{d \times d}$, then the matrix L_i has exactly the same block structure as D_i $(i = 1, \ldots, d)$, by construction. In particular, matrices $L_1, \ldots, L_{d_1}$ are formed from $L_1^1, \ldots, L_{d_1}^1$ respectively, etc. We now have $\mathcal{A}^{reg} = \oplus_{\alpha=1}^t \mathcal{A}_\alpha^{reg}$. This completes the proof. □

Using the last two lemmas, we can readily prove the following theorem.

Theorem 2.3.11. *The regular *-representation of* $\mathcal{A} := \oplus_{i=1}^t t_i \odot \mathbb{C}^{n_i \times n_i}$ *is equivalent to* $\oplus_{i=1}^t n_i \odot \mathbb{C}^{n_i \times n_i}$.

The Wedderburn decomposition of $\mathcal{A}^{reg}$ therefore takes the form

$$Q^* \mathcal{A}^{reg} Q = \oplus_{i=1}^t n_i \odot \mathbb{C}^{n_i \times n_i}, \tag{2.21}$$

for some suitable unitary matrix Q.

Comparing (2.20) and (2.21), we may informally say that the t_i and n_i values are equal for all i in the Wedderburn decomposition of a regular *-representation. We will also observe this in the numerical examples in Chapter 3.

2.4 Positive semidefinite matrices

This section is dedicated to the results for positive semidefinite matrices that we need in this thesis.

Recall that a complex matrix A is called Hermitian if $A^* = A$. Moreover, the eigenvalues of a Hermitian matrix are real. In the case where the matrices have real entries we can talk about symmetric matrices.

If $A, B \in \mathbb{C}^{n \times n}$ we can define the following inner product:

$$\langle A, B \rangle := \operatorname{trace}(AB^*) = \sum_{i,j=1}^n A_{ij} \overline{B}_{ij},$$

where $\overline{B}_{ij}$ is the complex conjugate of B_{ij}. Recall that we denoted the real symmetric $n \times n$ matrices by $\mathbb{S}^{n\times n}$. If $A, B \in \mathbb{S}^{n\times n}$ then

$$\langle A, B\rangle := \text{trace}(AB) = \sum_{i,j=1}^{n} A_{ij}B_{ij}.$$

This inner product induces the *Frobenius* (Euclidean) norm:

$$\|A\|^2 := \langle A, A\rangle = \text{trace}(AA^T) = \sum_{i,j=1}^{n} A_{ij}^2.$$

The following theorem presents equivalent characterizations of Hermitian positive semidefinite (psd) matrices.

Theorem 2.4.1 (see e.g., Section 7.2 in Horn and Johnson (1990))**.** *Let X be a Hermitian matrix. The following are equivalent:*

(1) $X \succeq 0$ (X is psd);

(2) $z^*Xz \geq 0\ \forall z \in \mathbb{C}^{n\times n}$;

(3) All eigenvalues of X are nonnegative;

(4) The determinants of all the principal minors of X are nonnegative;

(5) $X = LL^*$ for some $L \in \mathbb{C}^{n\times n}$.

A nonsingular matrix $X \succeq 0$ is called *positive definite* and we write $X \succ 0$. When the matrix has real entries, the vector z and the matrix L from Theorem 2.4.1 also have real entries and conjugation becomes transposition. An alternative notation is $\mathbb{S}_{+}^{n\times n}$ for positive semidefinite matrices and $\mathbb{S}_{++}^{n\times n}$ for positive definite matrices.

Lemma 2.4.2. *Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix. The following holds:*

$$A \succeq 0 \iff \begin{bmatrix} Re(A) & Im(A)^T \\ Im(A) & Re(A) \end{bmatrix} \succeq 0.$$

Proof. Recall that $A = \text{Re}(A) + \sqrt{-1}\text{Im}(A)$. $A \succeq 0$ so for any $z \in \mathbb{C}^n$, $z = a + \sqrt{-1}b$ with $a, b \in \mathbb{R}^n$ we have

$$[a + \sqrt{-1}b]^*(\text{Re}(A) + \sqrt{-1}\text{Im}(A))[a + \sqrt{-1}b] \geq 0.$$

This is equivalent to

$$[a^T - \sqrt{-1}b^T](\mathrm{Re}(A) + \sqrt{-1}\mathrm{Im}(A))[a + \sqrt{-1}b] \geq 0.$$

Thus,

$$\begin{aligned} & a^T\mathrm{Re}(A)a - a^T\mathrm{Im}(A)b + b^T\mathrm{Re}(A)b + b^T\mathrm{Im}(A)a \\ +\sqrt{-1}(a^T\mathrm{Re}(A)b + a^T\mathrm{Im}(A)a - b^T\mathrm{Re}(A)a + b^T\mathrm{Im}(A)b) & \geq 0. \end{aligned}$$

Since $-\mathrm{Im}(A) = \mathrm{Im}(A)^T$ (i.e., because $A = A^*$) we have $a^T\mathrm{Im}(A)a = -a^T\mathrm{Im}(A)a$, which yields $a^T\mathrm{Im}(A)a = 0$. Similarly $b^T\mathrm{Im}(A)b = 0$, and the other two terms from the imaginary part cancel each other.

Then the previous inequality is equivalent to

$$a^T\mathrm{Re}(A)a + a^T\mathrm{Im}(A)^Tb + b^T\mathrm{Re}(A)b + b^T\mathrm{Im}(A)a \geq 0,$$

which can be rewritten as

$$[a^T\ b^T]\begin{bmatrix} \mathrm{Re}(A) & \mathrm{Im}(A)^T \\ \mathrm{Im}(A) & \mathrm{Re}(A) \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} \geq 0.$$

Since this holds for any $\begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^{2n}$, we have

$$\begin{bmatrix} \mathrm{Re}(A) & \mathrm{Im}(A)^T \\ \mathrm{Im}(A) & \mathrm{Re}(A) \end{bmatrix} \succeq 0,$$

and this concludes the proof. $\square$

Any Hermitian matrix A has a *spectral decomposition*, that is:

$$A = \sum_{i=1}^{n} \lambda_i q_i q_i^* := Q\Lambda Q^*,$$

where q_i is the unit eigenvector corresponding to eigenvalue λ_i. Then $Q = [q_1, \dots q_n]$, $QQ^* = I$, and Λ is a diagonal matrix having $\Lambda_{ii} = \lambda_i$.

Any Hermitian matrix A has a *square root factorization.*

$$A^{\frac{1}{2}} := \sum_{i=1}^{n} \sqrt{\lambda_i} q_i q_i^*.$$

Note that $A^{\frac{1}{2}}A^{\frac{1}{2}} = A$. It follows that for any $X \succeq 0$

$$\mathrm{trace}(JX) = e^TXe = e^T(X^{\frac{1}{2}})^*X^{\frac{1}{2}}e = \|X^{\frac{1}{2}}e\|^2. \tag{2.22}$$

Definition 2.4.1. *The Schur complement of the invertible matrix D from the block matrix*

$$M := \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

is defined to be $S := A - BD^{-1}C$.

Lemma 2.4.3 (see e.g., Appendix A.5.5 in Boyd and Vandenberghe (2004))**.** *Let X be a symmetric matrix given by*

$$X := \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$$

and let S be the Schur complement of C in X. Then $X \succeq 0$ if and only if $A \succeq 0$ and $S \succeq 0$.

The following result is proved in the thesis of Gijswijt (2005) and, together with Theorem 2.4.1, it will turn out to be useful for proving that a matrix with a certain structure is positive semidefinite. For any $N \in \mathbb{S}^{n\times n}$, define:

$$M_N := \begin{pmatrix} 1 & \mathrm{diag}(N)^T \\ \mathrm{diag}(N) & N \end{pmatrix}. \tag{2.23}$$

Proposition 2.4.4 (see e.g., Proposition 7 in Gijswijt (2005))**.** *Let $N \in \mathbb{S}^{n\times n}$ be such that diag(N)=cNe for some $c \in \mathbb{R}$. Then the following are equivalent:*

(1) $M_N \succeq 0$;

(2) $N \succeq 0$ *and* $e^T Ne \geq (\mathrm{trace}(N))^2$.

2.5 Symmetry reduction of SDP instances

We consider the standard *primal* SDP problem:

$$\min_{X\succeq 0}\{\mathrm{trace}(A_0X) \mid \mathrm{trace}(A_kX) = b_k \ \ \forall\, k = 1,\dots,m\}, \tag{2.24}$$

where the Hermitian data matrices $A_i = A_i^* \in \mathbb{C}^{n\times n}$ $(i = 0,\dots,m)$ are linearly independent. The *dual problem* is formulated as follows:

$$\max_{y\in\mathbb{R}^m,\, S\succeq 0}\{b^Ty \mid \sum_{i=1}^{m} y_iA_i + S = A_0\}, \tag{2.25}$$

where y_i, $(i = 1, ..., m)$ are the dual variables, and the matrix S is also called the slack matrix.

We say that the SDP data matrices exhibit *algebraic symmetry* if the following assumption holds.

Assumption 1 (Algebraic symmetry). *There exists a matrix $\mathbb{C}*$-algebra, say $\mathcal{A}_{SDP}$ with dim$(\mathcal{A}_{SDP}) < n$, that contains the data matrices $A_0, \ldots, A_m$.*

Under this assumption, we may restrict the feasible set of problem (2.24) to its intersection with $\mathcal{A}_{SDP}$, as the following theorem shows (taken from De Klerk, Dobre, and Pasechnik (2010)). Related, but slightly less general, results are given by Gatermann and Parrilo (2004), De Klerk (2010), and others.

Theorem 2.5.1. *Let $\mathcal{A}_{SDP}$ denote a matrix $\mathbb{C}*$-algebra that contains the data matrices $A_0, \ldots, A_m$ of problem (2.24) as well as the identity. If problem (2.24) has an optimal solution, then it has an optimal solution in $\mathcal{A}_{SDP}$.*

Proof. By Theorem 2.1.5 we may assume that there exists a unitary matrix Q such that

$$Q^*\mathcal{A}_{SDP}Q = \oplus_{i=1}^{t} t_i \odot \mathbb{C}^{n_i \times n_i}, \tag{2.26}$$

for some integers t, n_i, and t_i $(i = 1, \ldots, t)$.

Since $A_0, \ldots, A_m \in \mathcal{A}$,

$$Q^*A_jQ =: \oplus_{i=1}^{t} t_i \odot A_j^{(i)} \quad (j = 0, \ldots, m)$$

for Hermitian matrices $A_j^{(i)} \in \mathbb{C}^{n_i \times n_i}$ where $i = 1, \ldots, t$ and $j = 0, \ldots, m$.

Now assume $\widetilde{X}$ is an optimal solution for (2.24). We have for each $i = 0, ..., m$:

$$\begin{aligned}
\text{trace}(A_j\widetilde{X}) &= \text{trace}(QQ^*A_jQQ^*\widetilde{X}) \\
&= \text{trace}((Q^*A_jQ)(Q^*\widetilde{X}Q)) \\
&= \text{trace} \oplus_{i=1}^{t} t_i \odot A_j^{(i)} Q^*\widetilde{X}Q \\
&=: \text{trace} \oplus_{i=1}^{t} t_i \odot A_j^{(i)} \bar{X},
\end{aligned}$$

where $\bar{X} := Q^*\widetilde{X}Q$.

The only elements of $\bar{X}$ that appear in the last expression are those in the diagonal blocks that correspond to the block structure of $Q^*\mathcal{A}Q$. We may therefore construct a matrix $\bar{X}' \succeq 0$ from $\bar{X}$ by setting those elements of $\bar{X}$ that are outside the blocks to zero, say

$$\bar{X}' = \oplus_{i=1}^{t} \left(\oplus_{k=1}^{t_i} \bar{X}_i^{(k)} \right)$$

where the $\bar{X}_i^{(k)} \in \mathbb{C}^{n_i \times n_i}$ $(k = 1, \ldots, t_i)$ are the diagonal blocks of $\bar{X}$ that correspond to the blocks of the i-th basic algebra, i.e., $t_i \odot \mathbb{C}^{n_i \times n_i}$ in (2.26). Thus, we obtain, for $j = 0, \ldots, m$,

$$\begin{aligned}
\operatorname{trace}(A_j \widetilde{X}) &= \operatorname{trace} \oplus_{i=1}^{t} t_i \odot A_j^{(i)} \bar{X}' \\
&= \sum_{i=1}^{t} \sum_{k=1}^{t_i} \operatorname{trace} \left(A_j^{(i)} \bar{X}_i^{(k)} \right) \\
&= \sum_{i=1}^{t} \operatorname{trace} \left(A_j^{(i)} \left[\sum_{k=1}^{t_i} \bar{X}_i^{(k)} \right] \right) \\
&=: \sum_{i=1}^{t} \operatorname{trace} \left((t_i \odot A_j^{(i)})(t_i \odot \bar{X}_i) \right),
\end{aligned} \tag{2.27}$$

where $\bar{X}_i = \frac{1}{t_i} \left[\sum_{k=1}^{t_i} \bar{X}_i^{(k)} \right] \succeq 0$. Defining

$$X := \sum_{i=1}^{t} t_i \odot \bar{X}_i$$

we have $X \succeq 0$, $X \in Q^* \mathcal{A}_{SDP} Q$ by (2.26), and

$$\operatorname{trace}(A_j \widetilde{X}) = \operatorname{trace}(Q^* A_j Q X) = \operatorname{trace}(A_j Q X Q^*), \quad (j = 0, \ldots, m),$$

by (2.27). Thus, $QXQ^* \in \mathcal{A}_{SDP}$ is an optimal solution of (2.24). □

In most applications, the data matrices $A_0, \ldots, A_m$ are real, symmetric matrices, and we may assume that $\mathcal{A}_{SDP}$ has a real basis (seen as a subspace of $\mathbb{C}^{n \times n}$). In this case, if (2.24) has an optimal solution, it has a real optimal solution in $\mathcal{A}_{SDP}$.

Corollary 2.5.2. *Assume the data matrices $A_0, \ldots, A_m$ in (2.24) are real symmetric. If $X \in \mathbb{C}^{n \times n}$ is an optimal solution of problem (2.24) then $Re(X)$ is also an optimal solution of this problem. Moreover, if $\mathcal{A}_{SDP}$ has a real basis, and $X \in \mathcal{A}_{SDP}$, then $Re(X) \in \mathcal{A}_{SDP}$.*

Proof. We have $\mathrm{trace}(A_kX) = \mathrm{trace}(A_k\mathrm{Re}(X))$ $(k = 0, \dots, m)$. Moreover, $X \succeq 0$ implies that $\mathrm{Re}(X) \succeq 0$. The second part of the result follows from the fact that, if $X \in \mathcal{A}_{SDP}$ and $\mathcal{A}_{SDP}$ has a real basis, then both $\mathrm{Re}(X) \in \mathcal{A}_{SDP}$ and $\mathrm{Im}(X) \in \mathcal{A}_{SDP}$. □

By Theorem 2.5.1, we may rewrite the SDP problem (2.24) as:

$$\min_{X \succeq 0} \left\{\mathrm{trace}(A_0X) \mid \mathrm{trace}(A_kX) = b_k \quad (k = 1, \dots, m),\ X \in \mathcal{A}_{SDP}\right\}. \tag{2.28}$$

Assume now that we have an orthogonal basis $B_1, \dots, B_d$ of $\mathcal{A}_{SDP}$. We set $X = \sum_{i=1}^d x_iB_i$ to get

$$\begin{aligned} & \min_{X \succeq 0} \left\{\mathrm{trace}(A_0X) \mid \mathrm{trace}(A_kX) = b_k \quad (k = 1, \dots, m), X \in \mathcal{A}_{SDP}\right\} \\ = & \min_{\sum_{i=1}^d x_iB_i \succeq 0} \left\{\sum_{i=1}^d x_i\mathrm{trace}(A_0B_i) \mid \sum_{i=1}^d x_i\mathrm{trace}(A_kB_i) = b_k, \right. \\ & \left. (k = 1, \dots, m)\right\}. \end{aligned} \tag{2.29}$$

If $\mathcal{A}_{SDP}$ has a basis that is a coherent configuration (see Definition 2.1.2), then we may assume that the B_i are zero-one matrices that sum to the all-ones matrix. In this case, adding the additional constraint $X \geq 0$ (i.e., X elementwise nonnegative) to problem (2.28) is equivalent to adding the additional constraint $x \geq 0$ to (2.29).

We may now replace the linear matrix inequality in the last SDP problem by an equivalent one,

$$\sum_{i=1}^d x_iB_i \succeq 0 \Longleftrightarrow \sum_{i=1}^d x_iQ^*B_iQ \succeq 0,$$

to get a block-diagonal structure, where Q is the unitary matrix that provides the Wedderburn decomposition of $\mathcal{A}_{SDP}$. In particular, we obtain

$$Q^*B_kQ =: \oplus_{i=1}^t t_i \odot B_k^{(i)}, \quad (k = 1, \dots, d)$$

for some Hermitian matrices $B_k^{(i)} \in \mathbb{C}^{n_i \times n_i}$ $(i = 1, \dots, t)$. Subsequently, we may delete any identical copies of blocks in the block structure to obtain a final reformulation. In particular, $\sum_{i=1}^d x_iQ^*B_iQ \succeq 0$ becomes

$$\sum_{k=1}^d x_k \oplus_{i=1}^t B_k^{(i)} \succeq 0.$$

Thus, we arrive at the final SDP reformulation:

$$\min_{x\in\mathbb{R}^d}\left\{\sum_{i=1}^{d} x_i \mathrm{trace}(A_0B_i) \mid \sum_{i=1}^{d} x_i \mathrm{trace}(A_kB_i) = b_k\ \forall k,\ \sum_{k=1}^{d} x_k(\oplus_{i=1}^{t} B_k^{(i)}) \succeq 0\right\}. \quad (2.30)$$

Note that the numbers $\mathrm{trace}(A_kB_i)$ ($k = 0,\ldots,m$, $i = 1,\ldots,d$) may be computed beforehand.

An alternative way to arrive at the final SDP formulation (2.30) is as follows. We may first replace the linear matrix inequality (LMI) $\sum_{i=1}^{d} x_iB_i \succeq 0$ by $\sum_{i=1}^{d} x_iL_i \succeq 0$, where the L_i ($i = 1,\ldots,d$) form the basis of the regular *-representation of $\mathcal{A}_{SDP}$. Now we may replace the new LMI using the Wedderburn decomposition (block-diagonalization) of the L_i's, and delete any duplicate blocks as before. These two approaches result in the same final SDP formulation, but the latter approach offers numerical advantages and is used to obtain the numerical results in Sections 3.5 and 3.6.

Note that, even if the data matrices A_i are real symmetric, the final block diagonal matrices in (2.30) may in principle be complex Hermitian matrices, since Q may be unitary (as opposed to real orthogonal). This poses no problem in theory, since interior point methods apply to SDP with Hermitian data matrices as well. If required, one may reformulate a Hermitian linear matrix inequality in terms of real matrices by applying Lemma 2.4.2 to each block in the LMI. Note that this doubles the size of the block.

Example 2.5.1. *Consider the problem of computing the ϑ' number of the pentagon, denoted C_5 (see McEliece, Rodemich, and Rumsey (1978) and Schrijver (1979)). If A is the adjacency matrix of C_5, we have:*

$$\vartheta'(C_5) = \max\{\mathrm{trace}(JX)\ :\ \mathrm{trace}(AX) = 0,\ \mathrm{trace}(X) = 1,\ X \succeq 0,\ X \geq 0\}. \quad (2.31)$$

The data matrices of this SDP are J, A, and I. Note that the data matrices are invariant under the action of the automorphism group of C_5 (i.e., the dihedral group on 5 elements).

In our notation, $\mathcal{A}_{SDP}$ is the centralizer ring of aut(C_5). Recall from Example 2.2.1 that this is the set of 5×5 symmetric circulant matrices. Moreover, the matrices B_i (i=1,2,3) mentioned there form a basis for the centralizer ring of aut(C_5) that satisfies the properties in Theorem 2.2.2.

Observing that we can assume $A = B_2$ we obtain the SDP:

$$
\begin{aligned}
\vartheta'(C_5) \quad = \quad & \max_{x\geq 0}\left\{\sum_{i=1}^{3} x_i \mathrm{trace}(JB_i) \;:\; \sum_{i=1}^{3} x_i \mathrm{trace}(B_2B_i)=0,\; \sum_{i=1}^{3} x_i \mathrm{trace}(B_i)=1,\right.\\
& \left.\sum_{i=1}^{3} x_i B_i \succeq 0\right\} = \max_{x\geq 0}\left\{5x_1+10(x_2+x_3) \;:\; 10x_2=0,\; 5x_1=1,\right.\\
& \left.\sum_{i=1}^{3} x_i B_i \succeq 0\right\} = \max_{x_3\geq 0}\left\{1+10x_3 \;:\; \frac{1}{5}B_1 + x_3B_3 \succeq 0\right\} = \max_{x_3\geq 0}\left\{1+10x_3 \;:\right.\\
& \left.\frac{1}{\sqrt{5}}D_1 + \sqrt{10}D_3 \succeq 0\right\}.
\end{aligned}
$$

*Via the regular *-representation we may replace the D_i's by the L_i's from Example 2.3.1 and obtain:*

$$
\begin{aligned}
\vartheta'(C_5) \quad = \quad & \max_{x_3\geq 0}(1+10x_3)\\
\text{s.t.} \quad & \frac{1}{5}I + \sqrt{10}x_3 \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{5}} \\ 0 & \frac{1}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{10}} & 0 \end{pmatrix} \succeq 0.
\end{aligned}
$$

Example 2.5.2. *We may obtain an even larger reduction via the Wedderburn decomposition of $\mathcal{A}_{SDP}$ mentioned in the previous example. Recall that we had:*

$$
\vartheta'(C_5) = \max_{x_3\geq 0}\left\{1+10x_3 \;:\; \frac{1}{5}B_1 + x_3B_3 \succeq 0\right\}.
$$

From Example 2.2.1 we know that $\{B_1 = I, B_2, B_3\}$ forms an association scheme, and from Section 2.2.1 we know that these matrices can be simultaneously diagonalized via the discrete Fourier transform matrix, say Q. We have $Q^*B_1Q = I$ and

$$
Q^*B_3Q = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & -\frac{2}{\sqrt{5}-1} & 0 & 0 & 0 \\ 0 & 0 & \frac{3-\sqrt{5}}{\sqrt{5}-1} & 0 & 0 \\ 0 & 0 & 0 & \frac{3-\sqrt{5}}{\sqrt{5}-1} & 0 \\ 0 & 0 & 0 & 0 & -\frac{2}{\sqrt{5}-1} \end{pmatrix}.
$$

Deleting repeated blocks, we obtain an LP in one variable:

$$
\begin{aligned}
\vartheta'(C_5) \quad = \quad & \max_{x_3\geq 0}(1+10x_3)\\
\text{s.t.} \quad & \frac{1}{5}I + x_3 \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{2}{\sqrt{5}-1} & 0 \\ 0 & 0 & \frac{3-\sqrt{5}}{\sqrt{5}-1} \end{pmatrix} \succeq 0.
\end{aligned}
$$

The optimal solution can be now computed by hand, yielding $x_3 = \frac{\sqrt{5}-1}{10}$*, hence* $\vartheta'(C_5) = \sqrt{5}$.

Chapter 3

Numerical block diagonalization of matrix $*$-algebras

Of particular interest in this chapter is the exploitation of a structure called algebraic symmetry, where the SDP data matrices are contained in a low-dimensional matrix $\mathbb{C}*$-algebra. Although this structure may seem exotic, it arises in a surprising number of applications, and first appeared in the papers by Schrijver (1979) on bounds for binary code sizes and McEliece, Rodemich, and Rumsey (1978) on bounds for the coclique number of a graph.

3.1 Introduction

The numerical block diagonalization of matrix $*$-algebras also has more recent applications. These are surveyed in the papers by De Klerk (2010), Gatermann and Parrilo (2004), and Vallentin (2009). The applications include bounds on kissing numbers, see Bachoc and Vallentin (2008); bounds on crossing numbers in graphs, see De Klerk, Maharry, Pasechnik, Richter, and Salazar (2006) and De Klerk, Pasechnik, and Schrijver (2007); bounds on code sizes, see Schrijver (2005), Gijswijt, Schrijver, and Tanaka (2006), and Laurent (2009); truss topology design, see Bai, De Klerk, Pasechnik, and Sotirov (2009) and Kanno, Ohsaki, Murota, and Katoh (2001); quadratic assignment problems, see De Klerk and Sotirov (2010a); and bounds on the chromatic number of a graph, see Gvozdenović and Laurent (2008a).

As we have seen in Section 2.1, matrix $\mathbb{C}*$-algebras have a canonical block diagonal structure after a suitable unitary transformation, so algebraic symmetry may be exploited. Block diagonal structure may in turn be exploited by interior point algorithms.

For some examples of SDP instances with algebraic symmetry, the required unitary transform can be explicitly computed, e.g., the work of Schrijver (2005) for the Terwilliger algebra. For other examples, see e.g., De Klerk, Maharry, Pasechnik, Richter, and Salazar (2006) and De Klerk, Pasechnik, and Schrijver (2007), it can not. In the latter case, we can perform numerical preprocessing to obtain the required unitary transformation. A suitable algorithm is given by Eberly and Giesbrecht (2004), but the focus there is on complexity and symbolic computation, as opposed to practical floating-point computation. Murota, Kanno, Kojima, and Kojima (2010) presented a practical randomized algorithm that can be used for the preprocessing of SDP instances with algebraic symmetry.

In this chapter, we propose another numerical preprocessing approach in the spirit of the work by Murota, Kanno, Kojima, and Kojima (2010), although the details are somewhat different (see Section 3.5). We demonstrate that the new approach may offer numerical advantages for certain group-symmetric SDP instances, in particular for the SDP instances from the paper of De Klerk, Pasechnik, and Schrijver (2007). We show how to solve a specific instance from this paper in a few minutes on a PC after preprocessing, whereas the original solution (as reported by De Klerk, Pasechnik, and Schrijver (2007)) required a week on a supercomputer. The reduction in computational time is possible because De Klerk, Pasechnik, and Schrijver (2007) use only partial symmetry reduction (i.e., regular *-representation) whereas our approach uses the full block-diagonalization of the regular *-representation. This technique is useful since the initial data set is too large to be handled by the method of Murota, Kanno, Kojima, and Kojima (2010).

3.2 Constructing the Wedderburn decomposition

Let $\mathcal{A} \subseteq \mathbb{C}^{n\times n}$ be a matrix $\mathbb{C}*$-algebra. To construct the Wedderburn decomposition from Theorem 2.1.5 we need to define the following subalgebra of $\mathcal{A}$:

$$\text{center}(\mathcal{A}) = \{X \in \mathcal{A} \mid XA = AX \text{ for all } A \in \mathcal{A}\}.$$

The center is a commutative subalgebra of $\mathcal{A}$ and according to Proposition 2.1.1 has a common set of orthonormal eigenvectors that we may view as the columns of a unitary matrix Q, i.e., $Q^*Q = I$. We arrange the columns of Q such that eigenvectors corresponding to the same eigenvalue are grouped together.

In what follows we assume that $\mathcal{A}$ contains the identity. According to Proposition 2.1.2 there exists a basis of orthogonal, self-adjoint idempotents of center($\mathcal{A}$). Let us denote this basis by $E_1, \ldots, E_t$.

The unitary transform $\mathcal{A}' := Q^*\mathcal{A}Q$ transforms the E_i matrices to zero-one diagonal matrices (say $E_i' := Q^*E_iQ$) that sum to the identity. We clearly have

$$\mathcal{A}' = \sum_{i=1}^{t} \mathcal{A}'E_i' = \sum_{i=1}^{t} \mathcal{A}'E_i'^2 = \sum_{i=1}^{t} E_i'\mathcal{A}'E_i'.$$

Each term $E_i'\mathcal{A}'E_i'$ ($i = 1, \ldots, t$) is clearly a matrix $\mathbb{C}*$-algebra. For a fixed i, the matrices in the algebra $E_i'\mathcal{A}'E_i'$ have a common nonzero diagonal block indexed by the positions of the ones on the diagonal of E_i'. Define $\mathcal{A}_i$ to be the restriction of $E_i'\mathcal{A}'E_i'$ to this principal submatrix. We now have

$$Q^*\mathcal{A}Q = \oplus_{i=1}^{t}\mathcal{A}_i. \tag{3.1}$$

We know from Proposition 2.1.3 that each $\mathcal{A}_i$ from (3.1) has to be a simple algebra. Moreover, each $\mathcal{A}_i$ contains the identity matrix.

Numerically, the decomposition (3.1) of $\mathcal{A}$ into simple $\mathbb{C}*$-algebras may be done using the following framework algorithm.

Algorithm 3 Decomposition of $\mathcal{A}$ into simple $\mathbb{C}*$-algebras

INPUT: A $\mathbb{C}*$-algebra $\mathcal{A}$ that contains I.
(i) Sample a generic element, say X, from center($\mathcal{A}$).
(ii) Perform the spectral decomposition of X to obtain a unitary matrix Q containing a set of orthonormal eigenvectors of X.
OUTPUT: A unitary matrix Q such that $Q^*\mathcal{A}Q$ gives the decomposition (3.1).

In step (i), we assume that a basis of center($\mathcal{A}$) is available. The generic element X is then obtained by taking a random linear combination of the basis elements. This approach is also used by Murota, Kanno, Kojima, and Kojima (2010), see e.g., Algorithm 1 there.

3.3 From simple to basic components

Further decomposition of each $\mathcal{A}_i$ in (3.1) is possible, according to Proposition 2.1.4. We may decompose each $\mathcal{A}_i$ in (3.1) as

$$U_i^*\mathcal{A}_iU_i = t_i \odot \mathbb{C}^{n_i \times n_i}$$

for some integers n_i and t_i and some unitary matrix U_i $(i = 1, \ldots, t)$. For the dimensions to agree, we must have

$$\sum_{i=1}^{t} n_i t_i = n \text{ and } \dim(\mathcal{A}) = \sum_{i=1}^{t} n_i^2,$$

since $Q^*\mathcal{A}Q = \oplus_{i=1}^{t}\mathcal{A}_i \subset \mathbb{C}^{n\times n}$.

Now let $\mathcal{B}$ denote a given basic matrix $*$-algebra over $\mathbb{C}$. We can compute the decomposition $U^*\mathcal{B}U = t \odot \mathbb{C}^{s\times s}$, say, where U is unitary and s and t are integers, as follows.

Algorithm 4 Decomposition of basic $\mathcal{B}$ into basic $\mathbb{C}*$-algebras

INPUT: A basic $\mathbb{C}*$-algebra $\mathcal{B}$.

(i) Sample a generic element, say B, from any maximal commutative matrix $\mathbb{C}*$-subalgebra of $\mathcal{B}$.

(ii) Perform a spectral decomposition of B, and let Q denote the unitary matrix of its eigenvectors.

(iii) Partition $Q^*\mathcal{B}Q$ into $t \times t$ square blocks, each of size $s \times s$, where s is the number of distinct eigenvalues of B.

(iv) Sample a generic element from $Q^*\mathcal{B}Q$, say B', and denote the ij-th block by B'_{ij}. We may assume that $B'_{11}, \ldots, B'_{1t}$ are unitary matrices (possibly after a suitable constant scaling).

(v) Define the unitary matrix $Q' := \oplus_{i=1}^{t} (B'_{1i})^*$ and replace $Q^*\mathcal{B}Q$ by $Q'^*Q^*\mathcal{B}QQ'$. Each block in the latter algebra equals $\mathbb{C}I_s$.

(vi) Permute rows and columns to obtain $P^TQ'^*Q^*\mathcal{B}QQ'P = t \odot \mathbb{C}^{s\times s}$, where P is a suitable permutation matrix.

OUTPUT: A unitary matrix $U := QQ'P$ such that $U^*\mathcal{B}U = t \odot \mathbb{C}^{s\times s}$.

A few remarks on Algorithm 4:

- Step (i) in Algorithm 4 may be performed by randomly sampling a generic element from $\mathcal{B}$.
- By the proof of Proposition 2.1.4, the diagonal blocks in step (iii) are the algebras $\mathbb{C}I_s$.
- By the proof of Proposition 2.1.4, the blocks $B'_{11}, \ldots, B'_{1t}$ used in step (v) are all unitary matrices (up to a constant scaling), so that Q' in step (v) is unitary too.

3.4 Symmetry reduction by sampling from the center of the regular *-representation

Let $\mathcal{A}_{SDP}$ be the matrix $\mathbb{C}*$-algebra that contains the data matrices $A_0, \ldots A_m$ of the following SDP:

$$\min_{x \in \mathbb{R}^d} \left\{ \sum_{i=1}^{d} x_i \mathrm{trace}(A_0 D_i) \mid \sum_{i=1}^{d} x_i \mathrm{trace}(A_k D_i) = b_k \ \forall k, \ \sum_{i=1}^{d} x_i D_i \succeq 0 \right\}. \tag{3.2}$$

Moreover, $D_1, \ldots, D_d$ is an orthonormal basis of $\mathcal{A}_{SDP}$.

Recall from Section 2.3 that we aim to construct the Wedderburn decomposition of the regular $*$-representation of $\mathcal{A}_{SDP}$, denoted $\mathcal{A}_{SDP}^{reg}$, whose basis is given by the matrices $L_1, \ldots L_d$. We showed in Section 2.3 that problem (3.2) is equivalent to

$$\min_{x \in \mathbb{R}^d} \left\{ \sum_{i=1}^{d} x_i \mathrm{trace}(A_0 D_i) \mid \sum_{i=1}^{d} x_i \mathrm{trace}(A_k D_i) = b_k \ \forall k, \ \sum_{i=1}^{d} x_i L_i \succeq 0 \right\}.$$

To compute the Wedderburn decomposition of $\mathcal{A}_{SDP}^{reg}$ we need to sample a generic element from center($\mathcal{A}_{SDP}^{reg}$) (see step (i) in Algorithm 3). To this end, assume $X := \sum_{k=1}^{d} x_k L_k$ is in the center of $\mathcal{A}_{SDP}^{reg}$. This is the same as assuming that for $j = 1, ..., d$:

$$\begin{aligned}
XL_j = L_j X \quad &\Leftrightarrow \quad \sum_{i=1}^{d} x_i L_j L_i = \sum_{i=1}^{d} x_i L_i L_j \\
&\Leftrightarrow \quad \sum_{i=1}^{d} x_i \sum_k (L_j)_{ki} L_k = \sum_{i=1}^{d} x_i \sum_k (L_i)_{kj} L_k \\
&\Leftrightarrow \quad \sum_k \left(\sum_{i=1}^{d} x_i (L_j)_{ki} - \sum_{i=1}^{d} x_i (L_i)_{kj} \right) L_k = 0 \\
&\Leftrightarrow \quad \sum_{i=1}^{d} x_i \left((L_j)_{ki} - (L_i)_{kj} \right) = 0 \quad \forall k = 1, \ldots, d,
\end{aligned}$$

where the last equality follows from the fact that the L_k form a basis for $\mathcal{A}_{SDP}^{reg}$.

To sample a generic element from center($\mathcal{A}_{SDP}^{reg}$) we may therefore proceed as outlined in Algorithm 5.

In the numerical computation we add the extra constraint $\sum_{i=1}^{d} x_i = 1$ to avoid the zero solution.

Algorithm 5 Obtaining a generic element of center($\mathcal{A}_{SDP}^{reg}$)

INPUT: A basis $L_1, \ldots, L_d$ of $\mathcal{A}_{SDP}^{reg}$.
(i) Compute a basis of the nullspace of the linear operator $\mathcal{L} : \mathbb{C}^d \to \mathbb{C}^{d^2}$ given by

$$\mathcal{L}(x) = \left[\sum_{i=1}^{d} x_i \left((L_j)_{ki} - (L_i)_{kj} \right) \right]_{j,k=1,\ldots,d}.$$

(ii) Take a random linear combination of the basis elements of the nullspace of $\mathcal{L}$ to obtain a generic element, say $\bar{x}$, in the nullspace of $\mathcal{L}$.
OUTPUT: $\bar{X} := \sum_{k=1}^{d} \bar{x}_k L_k$, a generic element of center($\mathcal{A}_{SDP}^{reg}$).

After obtaining the Wedderburn decomposition, we arrive at the final SDP reformulation:

$$\min_{x \in \mathbb{R}^d} \left\{ \sum_{i=1}^{d} x_i \mathrm{trace}(A_0 D_i) \ : \ \sum_{i=1}^{d} x_i \mathrm{trace}(A_k D_i) = b_k \ \forall k, \ \sum_{k=1}^{d} x_k \oplus_{i=1}^{t} L_k^{(i)} \succeq 0 \right\}. \quad (3.3)$$

Noticing that trace$(A_k D_i)$ ($k = 0, \ldots, m$ and $i = 1, \ldots, d$) are simply numbers that can be precomputed, we can now conclude this section with a summary of the symmetry reduction for problem (3.2).

Algorithm 6 Symmetry reduction of SDP (3.2)

INPUT: Data for SDP (3.2), and a real, orthonormal basis $D_1, \ldots, D_d$ of $\mathcal{A}_{SDP}$.
(i) Compute the basis $L_1, \ldots, L_d$ of $\mathcal{A}_{SDP}^{reg}$ as described in Section 2.3.
(ii) Obtain a generic element from center($\mathcal{A}_{SDP}^{reg}$) via Algorithm 5.
(iii) Decompose $\mathcal{A}_{SDP}^{reg}$ into simple $\mathbb{C}*$-algebras using Algorithm 3.
(iv) Decompose the simple $\mathbb{C}*$-algebras from step (iii) into basic $\mathbb{C}*$-algebras using Algorithm 4.
OUTPUT: The reduced SDP of the form (3.3).

3.5 Relation to an approach by Murota et al.

In this section we explain the relation between our approach and that of Murota, Kanno, Kojima, and Kojima (2010). These authors study matrix $*$-algebras over $\mathbb{R}$ (as opposed to $\mathbb{C}$). This is more complicated than studying matrix $\mathbb{C}*$-algebras, since there is no simple analogy of the Wedderburn decomposition theorem (Theorem 2.1.5) for matrix $*$-algebras over $\mathbb{R}$. While any simple matrix $\mathbb{C}*$-algebra is basic, there are three types of simple matrix $*$-algebras over $\mathbb{R}$; see the work of Murota, Kanno, Kojima, and Kojima (2010) for a detailed discussion.

For example, Algorithm 1 in their paper decomposes a basis of the circulant matrices into real blocks of size 1×1 or 2×2, while our Algorithm 3 will produce a complex diagonalization, i.e., 1×1 blocks that appear as complex conjugate pairs. Theorem 3.5.2 shows how the methods are related.

We assume now, as in Murota, Kanno, Kojima, and Kojima (2010), that $\mathcal{A}_{SDP}$ is a matrix $*$-algebra over $\mathbb{R}$, that

$$\mathcal{A}_{SDP} \cap \mathbb{S}^{n\times n} = \operatorname{span}\{A_0, \ldots, A_m\},$$

and that

$$\mathcal{A}_{SDP} = \langle \{A_0, \ldots, A_m\} \rangle.$$

In words, the A_i's generate $\mathcal{A}_{SDP}$ and form a basis for the symmetric part of $\mathcal{A}_{SDP}$.

Murota *et al.* (see Algorithm 1 in the paper of Murota, Kanno, Kojima, and Kojima (2010)) decompose $\mathcal{A}_{SDP}$ into simple components as follows:

Algorithm 1 in Murota, Kanno, Kojima, and Kojima (2010)

1. Choose a random $r = [r_0, \ldots, r_m]^T \in \mathbb{R}^{m+1}$.
2. Let $A := \sum_{i=0}^{m} r_i A_i$.
3. Perform the spectral decomposition of A to obtain an orthogonal matrix, say Q, of eigenvectors.
4. Make a k-partition of the columns of Q that defines matrices Q_i $(i = 1, \ldots, k)$ so that

$$Q_i^T A_p Q_j = 0 \quad \forall\, p = 0, \ldots, m, \quad i \neq j. \tag{3.4}$$

Similarly to our Algorithm 3, this algorithm involves sampling from the center of $\mathcal{A}_{SDP}$. To prove this, we will require the following lemma.

Lemma 3.5.1. *Assume $A = A^T \in \mathcal{A}_{SDP}$ has spectral decomposition $A = \sum_i \lambda_i q_i q_i^T$. Then, for any (eigen)values $r_i \in \mathbb{R}$, the matrix*

$$R := \sum_i r_i q_i q_i^T$$

is also in $\mathcal{A}_{SDP}$, provided that $r_i = r_{i'}$ whenever $\lambda_i = \lambda_{i'}$.

Proof. Let k denote the number of positive, distinct eigenvalues of A. Our goal is to show that the linear system

$$R = \alpha_1 A_1 + \ldots + \alpha_k A^k \tag{3.5}$$

has a (unique) solution, which implies that $R \in \mathcal{A}_{SDP}$. After diagonalization, since $r_i = r_{i'}$ whenever $\lambda_i = \lambda_{i'}$, the linear system (3.5) reduces to

$$r_i = \lambda_i^1 \alpha_1 + \ldots + \lambda_i^k \alpha_k \quad (i = 1, \ldots, k).$$

This is a linear system with variables α_i whose matrix of coefficients is a nonsingular Vandermonde matrix. Therefore, the values α_i are determined uniquely. Hence, $R \in \mathcal{A}_{SDP}$. □

Theorem 3.5.2. *The matrices $Q_i Q_i^T$ $(i = 1, \ldots, k)$ are symmetric, central idempotents of $\mathcal{A}_{SDP}$.*

Proof. For each i, let $E_i := Q_i Q_i^T$, and note that $E_i^2 = E_i$, i.e., E_i is idempotent. Also note that

$$\sum_{i=1}^{k} E_i = QQ^T = I. \tag{3.6}$$

Fix $p \in \{0, \ldots, m\}$. Then

$$\begin{aligned} E_i A_p E_j &= Q_i \left(Q_i^T A_p Q_j \right) Q_j^T \\ &= 0 \text{ if } i \neq j. \end{aligned}$$

By (3.6) we have

$$\sum_{i=1}^{k} E_i A_p = A_p \text{ and } \sum_{i=1}^{k} A_p E_i = A_p,$$

which implies that $E_j A_p E_j = A_p E_j$ and $E_j A_p E_j = E_j A_p$ respectively $(j = 1, \ldots, k)$. Thus, $E_j A_p = A_p E_j$ for all $j = 1, \ldots, k$. Since the A_i $(i = 1, \ldots, m)$ are generators of $\mathcal{A}_{SDP}$, this means that the E_j's $(j = 1, \ldots, k)$ are in the commutant of $\mathcal{A}_{SDP}$.

It remains to show that $E_j \in \mathcal{A}_{SDP}$ $(j = 1, \ldots, k)$. This follows directly from Lemma 3.5.1. Note that E_j and A share the set Q of eigenvectors, so $E_j \in \mathcal{A}_{SDP}$ $(j = 1, \ldots, k)$ by the lemma. Thus, E_j $(j = 1, \ldots, k)$ is in the center of $\mathcal{A}_{SDP}$.

□

Note that the matrix Q is implicitly used to construct the matrices E_j. In particular, the the k-partition of Q yields the matrices Q_j and then $E_j := Q_j Q_j^T$ are central idempotents, but they do not necessarily form a basis of the center of the matrix $*$-algebra.

3.6 Example: Bounds for crossing numbers

We know from Section 1.2.3 that the crossing number $\mathrm{cr}(G)$ of a graph G is the minimum number of intersections of edges in a drawing of G in the plane.

3.6.1 Symmetry reduction of the SDP relaxation

De Klerk, Maharry, Pasechnik, Richter, and Salazar (2006) showed that one may obtain a lower bound on $\mathrm{cr}(K_{r,s})$ via the optimal value of a suitable SDP, namely:

$$\mathrm{cr}(K_{r,s}) \geq \frac{s}{2}\left(s \min_{X \geq 0,\ X \succeq 0} \{\mathrm{trace}(MX) \ \mid \ \mathrm{trace}(JX) = 1\} - \left\lfloor \frac{r}{2} \right\rfloor \left\lfloor \frac{r-1}{2} \right\rfloor\right),$$

where M is a certain (given) matrix of order $n = (r-1)!$, and J is the all-ones matrix of the same size. The rows and columns of M are indexed by all the cyclic orderings of r elements. For this SDP problem the algebra $\mathcal{A}_{SDP}$ is a coherent configuration and an orthogonal (orthonormal) basis $B_1, \ldots, B_d$ $(D_1, \ldots, D_d)$ of $\mathcal{A}_{SDP}$ is available.

Some information on $\mathcal{A}_{SDP}$ is given in Table 3.1.

r	$n = (r-1)!$	$d := \dim(\mathcal{A}_{SDP})$	$d_1 := \dim(\mathcal{A}_{SDP} \cap \mathbb{S}^{n \times n})$
7	720	78	56
8	5040	380	239
9	40320	2438	1366

Table 3.1: Information on $\mathcal{A}_{SDP}$ for the crossing-number SDP instances

The instance corresponding to $r = 7$ was first solved in the paper of De Klerk, Maharry, Pasechnik, Richter, and Salazar (2006), by solving the partially reduced SDP (2.28), that in this case takes the form:

$$\min_{x \geq 0} \left\{ \sum_{i=1}^{d} x_i \mathrm{trace}(MB_i) \mid \sum_{i=1}^{d} x_i \mathrm{trace}(JB_i) = 1,\ \sum_{i=1}^{d} x_i B_i \succeq 0 \right\}.$$

The larger instances where $r = 8, 9$ were solved by De Klerk, Pasechnik, and Schrijver (2007) by solving the equivalent, but smaller problem:

$$\min_{x \geq 0} \left\{ \sum_{i=1}^{d} x_i \mathrm{trace}(MD_i) \mid \sum_{i=1}^{d} x_i \mathrm{trace}(JD_i) = 1, \ \sum_{i=1}^{d} x_i L_i \succeq 0 \right\}, \tag{3.7}$$

where the L_i's ($i = 1, \ldots, d$) form the basis of $\mathcal{A}_{SDP}^{reg}$.

In what follows we further reduce the latter problem by computing the Wedderburn decomposition of $\mathcal{A}_{SDP}^{reg}$ using Algorithm 6. We computed the basis of $\mathcal{A}_{SDP}^{reg}$ using a customized extension of the computational algebra package GRAPE developed by Soicher (2006), that in turn is part of the GAP (2008) routine library.

The Wedderburn decomposition results in block diagonalization of the L_i's ($i = 1, \ldots, d$), and the sizes of the resulting blocks are shown in Table 3.2.

r	$t_i = n_i$
7	3 (6×), 2 (4×), 1 (8×)
8	7 (2×), 5 (2×), 4 (9×), 3 (7×), 2 (4×), 1 (9×)
9	12 (8×), 11 (2×), 9 (6×), 7 (3×), 6 (5×), 5 (2×), 4 (2×), 3 (16×), 1 (5×)

Table 3.2: Block sizes in the decomposition $Q^* \mathcal{A}_{SDP}^{reg} Q = \oplus_i t_i \odot \mathbb{C}^{n_i \times n_i}$. Since $\mathcal{A}_{SDP}^{reg}$ is the regular $*$-representation of $\mathcal{A}_{SDP}$ we have $t_i = n_i$ for all i (see Theorem 2.3.11).

The difference between the sparsity patterns of a generic matrix in $\mathcal{A}_{SDP}^{reg}$ before and after symmetry reduction is illustrated in Fig. 3.1 when $r = 9$. In this case, $\mathcal{A}_{SDP}^{reg} \subset \mathbb{C}^{2438 \times 2438}$. Before symmetry reduction, there is no visible sparsity pattern. After $\mathcal{A}_{SDP}^{reg}$ is decomposed into simple components, a block-diagonal structure is visible, with largest block size 144. After the simple components are decomposed into basic components, the largest block size is 12.

After the Wedderburn decomposition, problem (3.7) becomes:

$$\min_{x \geq 0} \left\{ \sum_{i=1}^{d} x_i \mathrm{trace}(MD_i) \mid \sum_{i=1}^{d} x_i \mathrm{trace}(JD_i) = 1, \ \sum_{i=1}^{d} x_i \oplus_{k=1}^{t} L_i^{(k)} \succeq 0 \right\}. \tag{3.8}$$

3.6.2 Numerical results

Unless otherwise indicated, all computation was done using the SDPT3 solver developed by Toh, Todd, and Tütüncü (1999), and we used a Pentium IV PC with 2 GB

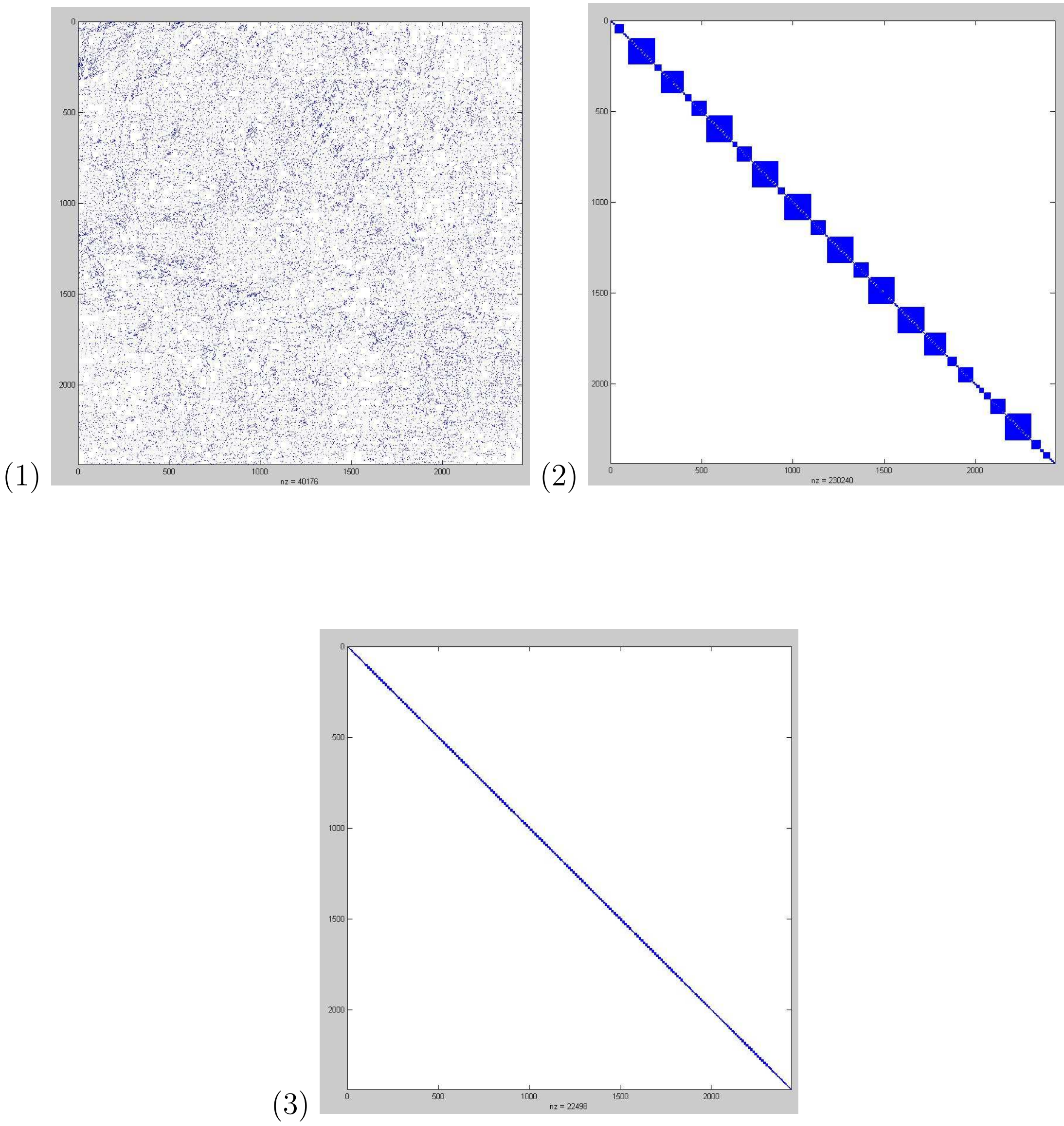

Figure 3.1: Sparsity pattern of $\mathcal{A}_{SDP}^{reg}$ for $r = 9$: (1) before any preprocessing, (2) after decomposition into simple $\mathbb{C}*$-algebras, and (3) after decomposition into basic $\mathbb{C}*$-algebras.

of RAM. Because of the huge dimension of the data (e.g., the text file containing information on matrices $L_1, \ldots L_d$ for $r = 9$ requires 1.3 GB of memory) we used the *sparse SDPA format*. The use of this format required some preprocessing for the solution of the relaxation (3.8). In this section we will present the necessary preprocessing for the SDP instances (3.7) and (3.8), i.e., before and after the numerical symmetry reduction (Algorithm 6). The results are presented in Table 3.3.

The first issue was that Algorithm 5 uses a basis of $\mathcal{A}_{SDP}^{reg}$, whereas the SDPA format can support only a basis of $\mathcal{A}_{SDP}^{reg} \cap \mathbb{S}^{d\times d}$.

Let $d_1 := \dim(\mathcal{A}_{SDP} \cap \mathbb{S}^{n\times n}) = \dim(\mathcal{A}_{SDP}^{reg} \cap \mathbb{S}^{d\times d})$. Recall that $\{B_1, \ldots, B_d\}$ forms a coherent configuration, so $B_i^T \in \mathcal{A}_{SDP}$ for each i. This last property is inherited by matrices D_i, L_i, and $L_i^{(k)}$, so we can group any nonsymmetric matrix $L_i^{(k)}$ with its transpose, say $L_{i^*}^{(k)}$. Moreover, since $\sum_{i=1}^{d} x_i \oplus_{k=1}^{t} L_i^{(k)} \in \mathbb{S}_{+}^{d\times d}$ we have $x_i = x_{i^*}$.

The same grouping was used to compute the coefficients $c_i := \mathrm{trace}(MD_i)$ and $e_i := \mathrm{trace}(JD_i)$ for $i = 1, \ldots d_1$. Hence, problem (3.8) is equivalent to:

$$\min_{x\geq 0} \left\{ \sum_{i=1}^{d_1} c_i x_i \ : \ \sum_{i=1}^{d_1} e_i x_i = 1, \ \sum_{i=1}^{d_1} x_i \oplus_{k=1}^{t} L_i^{(k)} \succeq 0 \right\}, \tag{3.9}$$

where $L_i^{(k)}$, $(i = 1, \ldots d_1)$ form a basis of $\mathcal{A}_{SDP}^{reg} \cap \mathbb{S}^{d\times d}$.

The SDPA format is a description of the following SDP:

$$\begin{array}{ll} \min\limits_{x\in\mathbb{R}^{d_1}} & c^T x \\ \text{s.t.} & \sum_{i=1}^{d_1} x_i F_i - F_0 \succeq 0, \end{array}$$

where the F_i (i=0,...,d_1) denote some real symmetric matrices and c is a vector of real numbers; these are the input data.

Since we minimize and $x_i \geq 0$ for all $i = 1, \ldots, d_1$, the constraint $\sum_{i=1}^{d_1} e_i x_i = 1$ may be replaced by $\sum_{i=1}^{d_1} e_i x_i - 1 \geq 0$. Hence, we represent in the SDPA format the following equivalent SDP:

$$\min \left\{ \sum_{i=1}^{d_1} c_i x_i \ : \ \sum_{i=1}^{d_1} x_i \oplus_{k=1}^{t} L_i^{(k)} \succeq 0, \ \sum_{i=1}^{d_1} e_i x_i - 1 \geq 0, \ x_i \geq 0, \ i = 1, ..., d_1 \right\}. \tag{3.10}$$

We define for $i = 1, \dots d_1$:

$$F_i = \begin{pmatrix} L_i^{(1)} & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & L_i^{(2)} & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & L_i^{(t)} & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & e_i & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

F_0 will have only one nonzero entry (equal to one) in the position of the value e_i from F_i.

For $r = 9$, the solution time reported by De Klerk, Pasechnik, and Schrijver (2007) was 7 days of wall-clock time on an SGI Altix supercomputer. With the numerical symmetry reduction this reduces to about 24 minutes on a Pentium IV PC, including the time for block diagonalization. For all three instances we obtained 6 digits of accuracy in the optimal value. Moreover, the same results were obtained for different random samples in the subroutines of Algorithms 5 and 4.

r	CPU time Algorithm 5	Solution time (3.8)	Solution time (3.7)
9	16 min 16 s	7 min 48 s	> 7 days†
8	4.7 s	3.2 s	5 min 22 s
7	0.04 s	0.6 s	2.7 s

Table 3.3: Solution times on a Pentium IV PC for the SDP instances before and after decomposition. † refers to wall-clock computation time on an SGI Altix supercomputer cluster.

The other dominant operation for the block diagonalization is the execution of Algorithm 5 (sampling from the center of $\mathcal{A}_{SDP}^{reg}$). The time required for this is shown in Table 3.3.

3.7 Example: Bounds for the ϑ'-number of graphs

Recall that an equivalent definition (due to De Klerk and Pasechnik (2002)) of the ϑ'-number of a graph G with adjacency matrix A is (see also Definition 2.31):

$$\vartheta'(G) := \max_{X \succeq 0,\, X \geq 0} \{\text{trace}(JX) \mid \text{trace}((A+I)X) = 1\}. \tag{3.11}$$

Note that the symmetry group $\mathcal{G}_{SDP}$ of this SDP coincides with the automorphism group of the graph. Thus, we may take $\mathcal{A}_{SDP}$ to be the centralizer ring of this group, as before.

De Klerk, Newman, Pasechnik, and Sotirov (2009) studied the ϑ' number of the so-called Erdős-Renyi graphs. These graphs, denoted $ER(q)$, are determined by a single parameter $q > 2$, which is prime. The number of vertices is $n = q^2 + q + 1$, but the dimension of $\mathcal{A}_{SDP}$ is only $2q + 11$. Note that, for example, if $q = 157$, then $n = 24807$, making it impossible to solve (3.11) without exploiting the symmetry.

The Wedderburn decomposition of $\mathcal{A}_{SDP}$ is not known in closed form, as explained by De Klerk, Newman, Pasechnik, and Sotirov (2009), and the numerical techniques proposed in this chapter may therefore be employed.

As before, we denote the zero-one basis of $\mathcal{A}_{SDP}$ by $B_1, \ldots, B_d$ and the basis of its regular *-representation $\mathcal{A}_{SDP}^{reg}$ by $L_1, \ldots, L_d$. De Klerk, Newman, Pasechnik, and Sotirov (2009) compute $\vartheta'(ER(q))$ for $q \leq 31$, by solving the SDP:

$$\vartheta'(ER(q)) = \min_{\sum_{i=1}^d x_i L_i \succeq 0, x \geq 0} \left\{ \sum_{i=1}^d x_i \mathrm{trace}(JB_i) \ : \ \sum_{i=1}^d x_i \mathrm{trace}((A+I)B_i) = 1 \right\}, \tag{3.12}$$

where A denotes the adjacency matrix of the graph $ER(q)$. The times required to compute the matrices $L_1, \ldots, L_d$ are given in Table 3.4 and were obtained using the GRAPE software, as before (see Soicher (2006)).

We can compute $\vartheta'(ER(q))$ for larger values of q by obtaining the Wedderburn decomposition of $\mathcal{A}_{SDP}^{reg}$ using Algorithm 6. The resulting block sizes are also given in Table 3.4. Note that the largest block size appearing in Table 3.4 is 3. As a result, all

q	d	Computing $L_1, \ldots, L_d$	$t_i = n_i$
157	325	1 h 22 min	2 (79 $\times$), 3 (1 $\times$)
101	213	21 min	2 (51 $\times$), 3 (1 $\times$)
59	129	4 min	2 (30 $\times$), 3 (1 $\times$)
41	93	1 min 22 s	2 (21 $\times$), 3 (1 $\times$)
31	73	37 s	2 (16 $\times$), 3 (1 $\times$)

Table 3.4: Block sizes in the decomposition $Q^* \mathcal{A}_{SDP}^{reg} Q = \oplus_i t_i \odot \mathbb{C}^{n_i \times n_i}$ where $\mathcal{A}_{SDP}^{reg}$ is the centralizer ring of $\mathrm{Aut}(ER(q))$.

the ϑ' values listed in the table were computed in a few seconds after the symmetry reduction; see Table 3.5.

To solve the SDP relaxation we used the same solver and data format as in Section 3.6, so all the preprocessing techniques described there were also applied to this problem.

In Table 3.5 the time required to perform the block diagonalization is shown in the second column (i.e., the execution time for Algorithm 5). In the last column we give the value of $\vartheta'(ER(q))$, since these values have not been computed previously for $q > 31$.

q	CPU time Algorithm 5	Solution time (3.12) after block diagonalization	Solution time (3.12)	$\vartheta'(ER(q))$
157	7.47	5.5	351.8	1834.394
101	1.34	1.4	70.3	933.137
59	0.21	0.8	11.6	408.548
41	0.047	0.61	6.5	233.389
31	0.018	0.49	3.4	151.702

Table 3.5: Time (s) to compute $\vartheta'(ER(q))$ with and without block diagonalization.

Note that for $q = 157$ the block diagonalization plus the solution of the resulting SDP takes a total of $7.47 + 5.5 \approx 13$ s, as opposed to the 351.8 s required to solve (3.12) without block diagonalization.

Chapter 4

Bounds for the symmetric circulant traveling salesman problem

This chapter is based on the work of De Klerk and Dobre (2009). Starting from a new SDP relaxation of the NP-complete traveling salesman problem (TSP) proposed by De Klerk, Pasechnik, and Sotirov (2008), we consider a special case where the SDP formulation can be reduced to a linear programming problem. Further, we compare, theoretically and numerically, the resulting bounds with the existing bounds in the literature.

4.1 Introduction

A (weighted) graph G is called *circulant* if its (weighted) adjacency matrix is circulant (see Section 2.2.1). Recall that we denoted the standard 0-1 basis of the symmetric circulant matrices by $\{B_0 := I, B_1, \ldots, B_d\}$, where $d := \lfloor n/2 \rfloor$, and n is the number of vertices in the graph. Thus,

$$(B_k)_{ij} := \begin{cases} 1 & \text{if } i - j = k \mod n \\ 0 & \text{otherwise} \end{cases} \qquad (k = 0, \ldots, d,\ i, j = 1, \ldots, n).$$

For circulant matrices it is usual to introduce some additional notation. If $\{t_1, \ldots, t_m\}$ is a subset of $\{0, 1, \ldots, d\}$, for some $m \leq d$, we define

$$C_n\langle t_1, \ldots, t_m\rangle := \sum_{i=1}^{m} B_{t_i}.$$

Thus, we will informally say that the circulant graph $C_n\langle t_1, \ldots, t_m\rangle$ consists of the stripes $t_1, \ldots, t_m$. In other words, we use the same notation for the circulant matrix $C_n\langle t_1, \ldots, t_m\rangle$ and the associated weighted circulant graph.

A natural question is whether a given combinatorial optimization problem becomes easier when restricted to circulant graphs. For example, the maximum clique and minimum graph coloring problems remain NP-hard for circulant graphs, and cannot be approximated within a constant factor, unless P=NP (see e.g., Codenotti, Gerace, and Vigna (1988)). Whether or not the Hamiltonian directed circuit problem restricted to directed circulant graphs remains NP-hard is still an open question; see Yang, Burkard, Cela, and Wöginger (1997), Heuberger (2002), and Bogdanowicz (2005). On the other hand, the shortest Hamiltonian path problem is polynomial solvable for undirected circulant graphs, as shown by Burkard and Sandholzer (1991). Likewise, deciding whether a circulant graph is Hamiltonian may be done in polynomial time, see again Burkard and Sandholzer (1991).

The symmetric circulant traveling salesman problem (SCTSP) is the problem of finding a Hamiltonian circuit of minimum length in a weighted, undirected, circulant graph. As far as we know, the complexity of the SCTSP is still open (see e.g., Van der Veen (1992) and Cook, Cunningham, Pulleyblank, and Schrijver (1998)). The best-known approximation algorithm for SCTSP is a 2-approximation algorithm (see Gerace and Greco (2008) and Van der Veen (1992)). The *bottleneck* TSP is known to be polynomially solvable in the circulant case, see Burkard and Sandholzer (1991). The study of the circulant TSP is motivated by practical applications, such as reconfigurable network design, see Medova (1994), and minimizing wallpaper waste, see Garfinkel (1977).

The main purpose of this chapter is to compare four lower bounds that can be obtained in polynomial time for the SCTSP:

1. We introduce a new linear programming bound derived from an SDP relaxation of TSP due to De Klerk, Pasechnik, and Sotirov (2008).

2. The second lower bound is due to Dantzig, Fulkerson, and Johnson (1954). Its optimal value coincides with the LP bound of Held and Karp (1970) (see e.g., Theorem 21.34 in Korte and Vygen (2008)), and it is commonly known as the Held-Karp (HK) bound.

3. The third bound (VdV) is due to Van der Veen (1992) and was introduced for the SCTSP. It is given as a closed-form expression and may be computed in linear time.

4. The fourth bound is the well-known 1-tree (1T) bound for TSP (see, e.g., Section 7.3 in Cook, Cunningham, Pulleyblank, and Schrijver (1998)).

We will show how bounds 1, 2, and 4 above may be computed more simply for circulant graphs than for general TSP. Subsequently, we will perform theoretical and empirical comparisons of the bounds.

4.2 Lower bounds for SCTSP

In this section we discuss four lower bounds for SCTSP.

SDP/LP bound

Let $K_n(D)$ denote a complete undirected graph on n vertices, with edge lengths (also called *weights* or *costs*) $D_{ij} = D_{ji} > 0$, $(i, j = 1, \ldots, n)$. Here D is called the matrix of distances. The Hamiltonian circuit in $K_n(D)$ of minimum length is often called the *optimal tour*.

It is shown by De Klerk, Pasechnik, and Sotirov (2008) that the following SDP provides a lower bound on the length of an optimal tour:

$$
\begin{aligned}
LP_{new} := \min \quad & \frac{1}{2}\mathrm{trace}(DX^{(1)}) \\
\text{s.t.} \quad & X^{(k)} \geq 0, \ \ (k = 1, \ldots, d) \\
& \sum_{k=1}^{d} X^{(k)} = J - I, \\
& I + \sum_{k=1}^{d} \cos(\frac{2ki\pi}{n}) X^{(k)} \succeq 0, \ \ (i = 1, \ldots, d) \\
& X^{(k)} \in \mathbb{S}^{n\times n}, \ \ (k = 1, \ldots, d),
\end{aligned}
\tag{4.1}
$$

where $d = \lfloor \frac{n}{2} \rfloor$ is the diameter of $\mathcal{C}_n$ (i.e., standard circuit on n vertices) and J denotes the all-ones matrix. Note that this problem involves nonnegative matrix variables $X^{(1)}, \ldots, X^{(d)}$ of order n. We will see in Section 4.3 that, if D is circulant, SDP formulation (4.1) reduces to an LP.

Held-Karp (HK) bound

One of the best-known linear programming (LP) relaxations of the TSP is the Held-Karp bound, defined as follows:

$$\begin{aligned} HK := \min \quad & \frac{1}{2}\mathrm{trace}(DX) \\ \text{s.t.} \quad & Xe = 2e, \\ & \mathrm{diag}(X) = 0, \\ & 0 \le X \le J, \\ & \sum_{i \in \mathcal{I}, j \notin \mathcal{I}} X_{ij} \ge 2 \ \forall\, \emptyset \ne \mathcal{I} \subset \{1, \ldots, n\}, \end{aligned} \tag{4.2}$$

where e denotes the all-ones vector and J the all-ones matrix, as before. Notice that when X is the adjacency matrix of a tour then we get the optimal value of the TSP, which shows that (4.2) is indeed a relaxation of the TSP.

The last constraints are called *subtour elimination inequalities* and model the fact that a Hamiltonian cycle is 2-connected. There are $2^n - 2$ subtour elimination inequalities, but even so this problem may be solved in polynomial time using the ellipsoid method (see e.g., Section 58.5 in Schrijver (2003)).

We will show how to simplify LP formulation (4.2) to an equivalent, smaller LP when the distance matrix D is circulant. The following theorem will allow us to restrict the optimization of (4.2) to the symmetric circulant matrices.

Theorem 4.2.1. *Let $\mathcal{A}$ denote the centralizer ring of a permutation group $\mathcal{G}$ and let $D \in \mathcal{A}$. If we have an optimal solution, X, for problem (4.2) then there exists an optimal solution, say $Y \in \mathcal{A}$, of problem (4.2).*

Proof. The fact that $D \in \mathcal{A}$ implies that $P^T DP = D$ for all $P \in \mathcal{G}$. We will show that if X is optimal for (4.2) then $Y := R(X)$ is also optimal for (4.2). Recall that $R(X)$ is the image of X under the Reynolds operator (see Section 2.2).

Since $Pe = e$, $P^T e = e$, and $Xe = 2e$ we have:

$$\begin{aligned} R(X)e &= \frac{1}{|\mathcal{G}|}\sum_{P \in \mathcal{G}} PXP^T e = \frac{1}{|\mathcal{G}|}\sum_{P \in \mathcal{G}} PXe \\ &= \frac{1}{|\mathcal{G}|}\sum_{P \in \mathcal{G}} 2Pe = \frac{1}{|\mathcal{G}|}\sum_{P \in \mathcal{G}} 2e = 2e. \end{aligned}$$

Permuting rows and columns preserves the zero diagonal, so diag(X) $= 0$ implies that diag(R(X)) $= 0$. Moreover, $R(X)$ averages over the permuted entries of X so that $0 \leq R(X) \leq J$ whenever $0 \leq X \leq J$.

To show that $R(X)$ is feasible for (4.2) we still have to prove that $R(X)$ satisfies the subtour elimination constraints. First notice that if P is a permutation matrix then matrices X and PXP^T are the adjacency matrices of two isomorphic graphs. Thus, the minimum cut in the graph with X as adjacency matrix equals the minimum cut in the graph with $R(X)$ as adjacency matrix. Thus,

$$\sum_{i \in \mathcal{I}, j \notin \mathcal{I}} (PXP^T)_{ij} \geq 2 \ \forall \emptyset \neq \mathcal{I} \subset \{1, \ldots, n\}.$$

Summing over all $P \in \mathcal{G}$ yields:

$$\sum_{P \in \mathcal{G}} \sum_{i \in \mathcal{I}, j \notin \mathcal{I}} (PXP^T)_{ij} \geq 2|\mathcal{G}| \ \forall \emptyset \neq \mathcal{I} \subset \{1, \ldots, n\}.$$

Thus,

$$\sum_{i \in \mathcal{I}, j \notin \mathcal{I}} (R(X))_{ij} \geq 2 \ \forall \emptyset \neq \mathcal{I} \subset \{1, \ldots, n\},$$

and $R(X)$ is therefore feasible for (4.2). Moreover, $R(X)$ is optimal since

$$\mathrm{trace}(DR(X)) = \mathrm{trace}(R(D)X) = \mathrm{trace}(DX)$$

by (2.16), and this concludes the proof of the theorem. $\square$

Recall that, for the SCTSP, the permutation group $\mathcal{G}$ is the dihedral group, and its centralizer ring is the set of symmetric circulant matrices. By Theorem 4.2.1, we can restrict the feasible set of (4.2) to the symmetric circulant matrices whose basis is $\{I = B_0, B_1, \ldots, B_d\}$. Since matrix D has zero on the diagonal we can ignore B_0 and write:

$$X := \sum_{p=1}^{d} x_p B_p \text{ and } D := \sum_{p=1}^{d} d_p B_p.$$

The objective in (4.2) reduces to:

$$\min \sum_{p=1}^{d} n d_p x_p,$$

if n is odd. If n is even, the last term becomes $\frac{1}{2} n d_d x_d$ instead of $n d_d x_d$.

To rewrite the subtour elimination constraints we will make use of a $\{0,1\}$ matrix denoted by $E_{\mathcal{I}}$. This matrix will have 1 in positions (i,j) and (j,i) if $i \in \mathcal{I}$, $j \notin \mathcal{I}$ and zeros elsewhere. Notice that

$$\frac{1}{2}\mathrm{trace}(E_{\mathcal{I}}X) = \sum_{i\in\mathcal{I}, j\notin\mathcal{I}} X_{ij}.$$

The subtour elimination constraints from (4.2) are then equivalent to:

$$\frac{1}{2}\sum_{p=1}^{d} x_p \mathrm{trace}(E_{\mathcal{I}}B_p) \geq 2 \ \forall\, \emptyset \neq \mathcal{I} \subset \{1,\ldots,n\}.$$

Notice that $\mathrm{diag}(X) = 0$ is implicit because $x_0 = 0$. Moreover, because $0 \leq X \leq J$ we have $0 \leq x_p \leq 1$, $p = 1,\ldots,d$. We have to split the constraint $Xe = 2e$ into two cases:

- For n odd: $Xe = 2e \Leftrightarrow \sum_{p=1}^{d} x_p B_p e = 2e \Leftrightarrow \sum_{p=1}^{d} x_p = 1$.
- For n even: $Xe = 2e \Leftrightarrow x_d B_d e + \sum_{p=1}^{d-1} x_p B_p e = 2e \Leftrightarrow \frac{1}{2}x_d + \sum_{p=1}^{d-1} x_p = 1$.

We can now give the simplified equivalent form of (4.2). For odd n, we have:

$$\begin{aligned} HK = \min \quad & \sum_{p=1}^{d} n d_p x_p \\ \text{s.t.} \quad & \sum_{p=1}^{d} x_p = 1, \\ & x_p \geq 0, \ \ (p = 1,\ldots,d) \\ & \frac{1}{2}\sum_{p=1}^{d} x_p \mathrm{trace}(E_{\mathcal{I}}B_p) \geq 2 \ \forall\, \emptyset \neq \mathcal{I} \subset \{1,\ldots,n\}. \end{aligned} \tag{4.3}$$

For even n, the last term in the objective function becomes $\frac{1}{2}n d_d x_d$, and the first constraint should be replaced by $\frac{1}{2}x_d + \sum_{p=1}^{d-1} x_p = 1$.

Van der Veen (VdV) bound

Let $D \in \mathbb{R}^{n\times n}$ be a symmetric circulant matrix and let $r = (r_0, r_1, \ldots, r_{\lfloor \frac{n}{2} \rfloor})$ be the vector that completely determines the entries of D (i.e., the first $d+1$ components on the first row). Recall that $\lfloor \frac{n}{2} \rfloor$=d.

Assume now that $r_0 = 0$ (which is the case for the TSP) and assume that the r_i are distinct. Define a permutation Φ such that $\Phi(0) = 0$ and Φ sorts the values of r in ascending order. Let $\gcd(t_1, \ldots, t_m)$ denote the greatest common divisor of given natural numbers $t_1, \ldots, t_m$. A necessary and sufficient condition for Hamiltonicity of a circulant graph is given by the following theorem.

Theorem 4.2.2 (Burkard and Sandholzer (1991))**.** *The circulant graph $C_n\langle t_1, \ldots, t_m\rangle$, with vertex set $\{0, 1, \ldots, n-1\}$, consists of $\gcd(n, t_1, \ldots, t_m)$ components ($m \in \{1, \ldots, \lfloor \frac{n}{2} \rfloor\}$). Each component is a graph on $\frac{n}{\gcd(n,t_1,\ldots,t_m)}$ vertices. The vertices in component α ($\alpha = 0, 1, \ldots, \gcd(n, t_1, \ldots, t_m) - 1$) are:*

$$\left\{ (\alpha + k \ \gcd(n, t_1, \ldots, t_m)) \ mod \ n \ \middle| \ k \in 0, 1, \ldots, \frac{n}{\gcd(n, t_1, \ldots, t_m)} - 1 \right\}. \quad (4.4)$$

Moreover, $C_n\langle t_1, \ldots, t_m\rangle$ is Hamiltonian if and only if $\gcd(n, t_1, \ldots, t_m)$=1.

Let l be the smallest integer such that $\gcd(n, \Phi(1), \ldots, \Phi(l)) = 1$. Van der Veen (1992) shows that one can construct a Hamiltonian tour using edges only from stripes $\Phi(1), \ldots, \Phi(l)$. Following his notation, we define:

$$\mathcal{GCD}(\Phi(k)) := \gcd(\mathcal{GCD}(\Phi(k-1)), \Phi(k)), \quad k = 1, \ldots, d, \quad (4.5)$$

and $\mathcal{GCD}(\Phi(0)) := n$. Further, we can assume without loss of generality (see Van der Veen (1992)) that

$$n = \mathcal{GCD}(\Phi(0)) > \mathcal{GCD}(\Phi(1)) > \ldots > \mathcal{GCD}(\Phi(l)) = 1. \quad (4.6)$$

Then Theorem 7.4.2 from Van der Veen (1992) shows that the following value is a lower bound for the SCTSP:

$$VdV := \sum_{i=1}^{l} \{(\mathcal{GCD}(\Phi(i-1)) - \mathcal{GCD}(\Phi(i)))r_{\Phi(i)}\} + r_{\Phi(l)}. \quad (4.7)$$

The term $\sum_{i=1}^{l} \{(\mathcal{GCD}(\Phi(i-1)) - \mathcal{GCD}(\Phi(i)))r_{\Phi(i)}\}$ gives the weight of a shortest Hamiltonian path obtained via the nearest neighbor rule. The last term reflects the fact that each Hamiltonian cycle must include an edge of weight at least $r_{\Phi(l)}$.

1-tree (1T) bound

Another famous lower bound for TSP is the minimum-cost 1-tree bound.

Definition 4.2.1. *Let G=(V,E) denote an undirected graph with edge costs c_e, for each $e \in E$, and let $v_1 \in V$. Two edges incident with node v_1 plus a spanning tree of $G \setminus \{v_1\}$ is called a 1-tree in G.*

Definition 4.2.2. *Let $G = (V, E)$ denote an undirected graph with edge costs c_e, for each $e \in E$, and let $v_1 \in V$. Let $\delta(v_1)$ denote the set of edges incident to v_1. Let $A = min\{c_e + c_f \mid e, f \in \delta(v_1)\}$ and let B be the cost of a minimum spanning tree in $G \setminus \{v_1\}$. Then $A + B$ is a lower bound for the TSP on G, called a 1-tree bound.*

For circulant graphs, we can compute the 1-tree bound more simply than for general graphs, as we will show in Theorem 4.2.4. Recall that we can construct a minimum-cost spanning tree using the (greedy) Kruskal algorithm. This algorithm starts with an arbitrary edge of lowest cost, and recursively constructs a spanning tree by adding an edge of lowest possible cost to the current forest so that adding this edge does not form a cycle.

As a consequence of Theorem 4.2.2, after using all possible edges from the lowest-cost stripe, we may assume that the Kruskal algorithm has constructed $x := \mathcal{GCD}(\Phi(1))$ components (i.e., disjoint paths). Moreover, by (4.4) we can describe these disjoint paths as:

$$P_\alpha := \left\{(\alpha + k\,\Phi(1)) \mod n \mid k \in 0, 1, \ldots, \frac{n}{x} - 1\right\}, \quad \alpha = 0, \ldots, x-1. \tag{4.8}$$

An important observation for our purposes is that these paths cover all the vertices; any edge that is subsequently added by the Kruskal algorithm will therefore connect two of these paths.

Now we fix v. According to the construction above we have $v \in P_{v \bmod x}$. Under assumption (4.6), we have $(v+\Phi(i)) \mod n \in P_{(v+\Phi(i)) \mod x}$, for every $i = 2, \ldots, l$. Thus, for each i, the edge $\{v, (v + \Phi(i)) \mod n\}$ connects the paths $P_{v \mod x}$ and $P_{(v+\Phi(i)) \mod x}$.

Lemma 4.2.3. *For any $k = 0, \ldots, \frac{n}{\mathcal{GCD}(\Phi(1))}$ and for any $i = 2, \ldots, l$ and $v \in V$, the edge $\{(v+k\Phi(1)) \mod n, (v+k\Phi(1)+\Phi(i)) \mod n\}$ connects the paths $P_{v \mod x}$ and $P_{(v+\Phi(i)) \mod x}$.*

Proof. By (4.8), $(v+k\Phi(1)) \mod n$ belongs to $P_{v \mod x}$. For any $i \in \{2,\dots,l\}$ we have

$$\begin{aligned}(v+k\Phi(1)+\Phi(i)) \mod n &= ((v+\Phi(i))+k\Phi(1)) \mod n \\ &= ((v+\Phi(i)) \mod n + k\Phi(1)) \mod n.\end{aligned}$$

Since $(v+\Phi(i)) \mod n$ belongs to $P_{(v+\Phi(i)) \mod x}$, using (4.8) again we have that $(v+k\Phi(1)+\Phi(i)) \mod n \in P_{(v+\Phi(i)) \mod x}$. □

The lemma shows that we can always connect two distinct paths P_{α_1} and P_{α_2} ($\alpha_1 \neq \alpha_2$) using an edge of cost $r_{\Phi(i)}$, for any $i = 2,\dots,l$, in more than one way. Now we can prove the following.

Theorem 4.2.4. *Let G be a circulant graph on n vertices. Let $\Phi(1)$ denote the stripe of minimum nonzero cost. The value of a minimum cost 1-tree equals the value of a minimum cost spanning tree plus the value of an edge of lowest cost whenever $\Phi(1) \neq \frac{n}{2}$. If n is even and $\Phi(1) = \frac{n}{2}$, then the value of a minimum cost 1-tree equals the value of a minimum cost spanning tree plus the cost of an edge of second-lowest cost.*

Proof. We will assume $\gcd(n,\Phi(1)) \neq 1$, since the case $\gcd(n,\Phi(1)) = 1$ is trivial.

Fix $v_1 \in V$, and assume no two stripes have the same cost and $\Phi(1) \neq \frac{n}{2}$. Because of the circulant structure we have two edges of minimum cost with an endpoint at v_1. Start constructing a minimum spanning tree from v_1 using Kruskal's algorithm (denote the first added edge by e_t). After adding the edges of minimum cost Kruskal's algorithm has constructed $\gcd(n,\Phi(1))$ disjoint paths covering the vertices of G with edges of lowest cost. After this step any other edge of lowest cost added to the current forest will create a cycle. Let the path with an endpoint at v_1 be P_{v_1}, and denote the other endpoint of this path by v_2. Connect the paths obtained before using edges of other costs (again using Kruskal's algorithm), but do not connect P_{v_1} via v_1 (this is always possible according to Lemma 4.2.3). When the minimum spanning tree is constructed add the edge $e_{12} := v_1v_2$. Call the resulting structure T.

By construction v_1 has degree 2 in T. The edges that connect v_1 to T are e_{12} and e_t. Notice that both have lowest cost. Therefore, $e_t + e_{12}$ is minimum among the sum of the costs of two edges incident to v_1, which shows that T is a 1-tree. Since v_1 was arbitrarily chosen this concludes the first part of the proof. The second part is similar and is therefore omitted. □

4.3 A new linear programming bound for SCTSP

In this section we show how to reduce the SDP formulation in (4.1) to an equivalent LP whenever the distance matrix D is circulant. The following theorem will allow us to restrict the optimization of (4.1) to the symmetric circulant matrices in the case of the SCTSP.

Theorem 4.3.1. *Let $\mathcal{A}$ denote the centralizer ring of a permutation group $\mathcal{G}$ and let $D \in \mathcal{A}$. If we have an optimal solution, $X^{(1)}, \ldots, X^{(k)}$, for problem (4.1) then $\{R(X^{(1)}), \ldots, R(X^{(k)})\} \subset \mathcal{A}$ is also an optimal solution of (4.1), where R denotes the Reynolds operator of the group $\mathcal{G}$.*

Proof. Since $D \in \mathcal{A}$, D is invariant under the action of the permutation matrices $P \in \mathcal{G}$, that is, $P^T DP = D$ for all $P \in \mathcal{G}$. We will show that if $X^{(k)}$, $(k = 1, \ldots, n)$ are feasible for (4.1) then $Y^{(k)} := R(X^{(k)})$ are also feasible for (4.1). For simplicity of notation we will show this for a fixed k, but everything holds for any $k = 1, \ldots, d$.

If $X^{(k)} \geq 0$ and symmetric, then by permuting rows and columns and adding elements we again obtain a symmetric, positive matrix, so $R(X^{(k)}) \geq 0$ and $R(X^{(k)}) \in \mathcal{S}^n$.

$R(X^{(k)})$ is a linear mapping so $R(\sum_{k=1}^d X^{(k)}) = \sum_{k=1}^d R(X^{(k)})$ and $R(J - I) = R(J) - R(I)$. Notice that $R(J) = J$ and $R(I) = I$. We thus obtain $\sum_{k=1}^d R(X^{(k)}) = J - I$. Using $R(I) = I$ and the linearity of R, from

$$I + \sum_{k=1}^{d} \cos(\frac{2ki\pi}{n})X^{(k)} \succeq 0, \ \ (i = 1, \ldots, d)$$

we obtain

$$I + \sum_{k=1}^{d} \cos(\frac{2ki\pi}{n})R(X^{(k)}) \succeq 0, \ \ (i = 1, \ldots, d).$$

We have seen that $R(X^{(k)})$, $(k = 1, \ldots, d)$, are feasible. Furthermore,

$$\mathrm{trace}(DR(X^{(1)})) = \mathrm{trace}(R(D)X^{(1)}) = \mathrm{trace}(DX^{(1)})$$

by (2.16), and this concludes the proof of the theorem. $\square$

Recall that, for the SCTSP, the permutation group $\mathcal{G}$ is the dihedral group, and its centralizer ring is the set of symmetric circulant matrices. Now let us restrict the feasible set to the circulant matrices. For each $X^{(k)}$, $(k = 1, \ldots, d)$, we can write

$$X^{(k)} := \sum_{p=1}^{d} x_p^{(k)} B_p, \tag{4.9}$$

where $\{B_0 = I, B_1, \ldots, B_d\}$ forms the standard basis for the symmetric circulant matrices, as before.

The matrix of distances D has zeros on the diagonal, and the variables $x_0^{(k)}$ may therefore be set to zero. Since the B_i's are 0-1 matrices, $X^{(k)} \geq 0$ is equivalent to $x_p^{(k)} \geq 0$, $(k, p = 1, \ldots, d)$; and using (4.9) we obtain the equivalent form of (4.1):

$$\begin{aligned} LP_{new} = \min \quad & \frac{1}{2} \sum_{p=1}^{d} x_p^{(1)} \operatorname{trace}(DB_p) \\ \text{s.t.} \quad & x_p^{(k)} \geq 0, \quad (k, p = 1, \ldots, d) \\ & \sum_{k=1}^{d} \sum_{p=1}^{d} x_p^{(k)} B_p = J - I, \\ & I + \sum_{k=1}^{d} \sum_{p=1}^{d} \cos(\frac{2ki\pi}{n}) x_p^{(k)} B_p \succeq 0, \quad (i = 1, \ldots, d). \end{aligned} \tag{4.10}$$

Let Q denote the discrete Fourier transform matrix. Then we may diagonalize the basis matrices via $Q^* B_p Q = \Lambda^{(p)}$, where $\Lambda^{(p)} := \operatorname{diag}(\lambda_j^{(p)})$, $(j = 0, \ldots, n-1)$, is the diagonal matrix containing the eigenvalues of B_p.

We have

$$\lambda_j^{(p)} = 2\cos(\frac{2\pi jp}{n}) \;\; (p = 1, \ldots, d, \;\; j = 0, \ldots, n-1), \;\; \text{if } n \text{ is odd} \tag{4.11}$$

and

$$\lambda_j^{(p)} = 2\cos(\frac{2\pi jp}{n}) \;\; (p = 1, \ldots, d-1, \;\; j = 0, \ldots, n-1), \;\; \text{if } n \text{ is even} \tag{4.12}$$

$$\lambda_j^{(d)} = \cos(\frac{2\pi jd}{n}), \;\; (j = 0, \ldots, n-1), \;\; \text{if } n \text{ is even.} \tag{4.13}$$

Because of the simultaneous diagonalization of the B_i, (4.10) reduces to an LP, as we will now show.

Let us write

$$D = \sum_{i=1}^{d} d_i B_i. \tag{4.14}$$

Clearly

$$\mathrm{trace}(B_iB_j) = 0 \;\; if \;\; i \neq j.$$

Multiplying (4.14) by B_p to the right and taking into account that B_i and D are symmetric, using the previous relation we obtain

$$\mathrm{trace}(DB_p) = d_p \mathrm{trace}(B_p^2) = cd_p, \tag{4.15}$$

where $c = 2n$ for $p = 1, \ldots, d$. For n even we have an exception: $c = n$ when $p = d$.

We will now transform each linear matrix equality into n linear inequalities. To this end, note that $J - I = \sum_{p=1}^{d} B_p$. Then, using the diagonalization, the relation

$$\sum_{k=1}^{d} \sum_{p=1}^{d} x_p^{(k)} B_p = \sum_{p=1}^{d} B_p$$

reduces to

$$\sum_{k=1}^{d} \sum_{p=1}^{d} x_p^{(k)} \lambda_j^{(p)} = \sum_{p=1}^{d} \lambda_j^{(p)}, \;\; (j = 0, \ldots, n-1), \tag{4.16}$$

where the eigenvalues $\lambda_j^{(p)}$ are defined in (4.11), (4.12), and (4.13).

Finally, again using the diagonalization, the d linear matrix inequalities

$$I + \sum_{k=1}^{d} \sum_{p=1}^{d} \cos(\frac{2ki\pi}{n}) x_p^{(k)} B_p \succeq 0, \;\; (i = 1, \ldots, d)$$

reduce to the nd linear inequalities

$$1 + \sum_{k=1}^{d} \sum_{p=1}^{d} \lambda_j^{(p)} \cos(\frac{2ki\pi}{n}) x_p^{(k)} \geq 0, \;\; (i = 1, \ldots, d, \;\; j = 0, \ldots, n-1). \tag{4.17}$$

We can now state the LP reformulation of (4.10):

$$\begin{aligned} LP_{new} = \min \quad & \frac{1}{2} \sum_{p=1}^{d} cd_p x_p^{(1)} \\ \text{s.t.} \quad & x_p^{(k)} \geq 0, \quad (k, p = 1, \ldots, d) \\ & \sum_{k=1}^{d} \sum_{p=1}^{d} \lambda_j^{(p)} x_p^{(k)} = \sum_{p=1}^{d} \lambda_j^{(p)}, \quad (j = 0, \ldots, n-1) \qquad (4.18) \\ & 1 + \sum_{k=1}^{d} \sum_{p=1}^{d} \lambda_j^{(p)} \cos(\frac{2ki\pi}{n}) x_p^{(k)} \geq 0, \quad (i = 1, \ldots, d, \;\; j = 0, \ldots, n-1). \end{aligned}$$

4.4 Numerical results

In this section we present numerical results for the new SDP/LP bound and the other bounds stated in Section 4.2 (i.e., 1T bound, HK bound, and VdV bound); see Table 4.1. The matrices in Table 4.1 have dimensions between 6 and 81, and were generated in such a way as to avoid *trivial solutions*.

By *trivial solutions* we refer to instances of SCTSP that are polynomially solvable. The most obvious case is where the number of vertices n is prime. Another example is the case where $\mathcal{GCD}(\phi(1)) = 1$ is polynomially solvable. In both situations an optimal tour can be constructed using only edges from the stripe of lowest cost (i.e., $\phi(1)$).

Moreover, at least one of the two heuristics proposed by Van der Veen (1992) gives the optimal value of a tour (in polynomial time) in each of the following cases:

- $n = p^2$, where $p \geq 2$ is a prime number;
- $\mathcal{GCD}(\phi(l-1)) = 2$;
- $\mathcal{GCD}(\phi(l-1)) = d$ and both $\phi(l-1)$ and $\phi(l)$ are odd numbers;
- $l = 2$, $\frac{n}{\mathcal{GCD}(\phi(1))}$ is odd and $\mathcal{GCD}(\phi(1)) \geq \frac{n}{\mathcal{GCD}(\phi(1))} - 1$;
- $l = 2$, $\mathcal{GCD}(\phi(l-1)) \leq 6$ and $\frac{r_{\phi(l+1)} - r_{\phi(l)}}{r_{\phi(l)} - r_{\phi(l-1)}} \geq \mathcal{GCD}(\phi(l-1)) - 2$.

In these five cases the value of an optimal tour is attained under the assumption that the costs of the stripes are distinct.

The only polynomial-solvable instances in Table 4.1 are $Dt14$, $Dt15$, and $Dt16$ since they prove that there is no dominance between the SDP/LP bound and the $1T$ bound.

The LP problems were solved using the Matlab® toolbox Yalmip (see Löfberg (2004)) together with the optimization solver Sedumi (see Sturm (1999)). The optimal values of the SCTSP instances were computed using the Concorde[1] software for TSP. Because of the small sizes of the instances, all the values in the tables could be computed in a few seconds on a standard Pentium IV PC.

A few remarks on Table 4.1:

- The HK and VdV bounds coincide for all the instances in the table.

[1]The Concorde software is available at `http://www.tsp.gatech.edu/concorde/`

matrix	n	l	$\mathcal{GCD}(\phi(l-1))$	SDP/LP	1T	HK	VdV	optimum
D1	54	3	3	2,114	2,140	2,157	2,157	2,174
D10	39	3	3	547.868	550	552	552	553
D11	57	2	3	2,022.715	2,119	2,181	2,181	2,243
D17	36	2	4	4,877.80	4,902	4,916	4,916	4,944
Dt4	24	2	6	123.91	125	126	126	128
Dt6	24	2	6	2,448.08	3,095	3,690	3,690	3,694
Dt14	8	2	2	57.17	57	58	58	58
Dt15	8	3	3	58.34	58	60	60	60
Dt16	6	2	2	43.50	43	44	44	44
Dt18	64	2	4	25,583	26,901	27,484	27,484	27,538
Dtt1	81	4	3	1,316.75	1,590	1,680	1,680	1,680
Dtt2	63	2	7	2,188.52	3,375	3,696	3,696	3,930

Table 4.1: Numerical comparison of the four lower bounds from Section 4.2 for SCTP instances.

- The HK and VdV bounds are the best bounds in all cases but do not always equal the optimal value of the SCTSP instance in question.
- The new LP bound is always weaker than the HK and VdV bounds for the test problems and is even lower than the 1T bound for a few instances. Adding the subtour elimination inequalities to the new LP did not result in better bounds than HK for any of the instances in the table.

The instances from Table 4.1 are available online at:
`http://lyrawww.uvt.nl/~cdobre/SCTSP_instances.rar`.

4.5 Theoretical comparison of bounds

Based on the numerical results presented in the previous section, we may conjecture certain relations between the bounds, such as VdV = HK $\geq LP_{new}$. On the other hand, we have been able to prove only that VdV $\geq$ 1T (cf. Theorem 4.5.1) and HK $\geq$ VdV (cf. Theorem 4.5.3). It is also well known (see e.g., Cook, Cunningham, Pulleyblank, and Schrijver (1998)) that HK $\geq$ 1T. Thus, we have the sandwich-theorem result

$$1T \leq VdV \leq HK.$$

Theorem 4.5.1. *For SCTSP, the VdV bound is at least as good as the $1T$ bound.*

Proof. Recall that $\Phi(1)$ denotes the stripe of lowest cost. From (4.7) we have that VdV equals the length of a minimum-weight Hamiltonian path plus the weight of an edge of cost $r_{\Phi(l)}$. Moreover, the weight of a minimum Hamiltonian path is always greater than or equal to the weight of a minimum-weight spanning tree. The required result now follows from Theorem 4.2.4. $\square$

Thus, we have VdV $\geq$ 1T. Further, it was shown by De Klerk, Pasechnik, and Sotirov (2008) that, for general TSP, HK does not dominate the SDP bound in (4.1) or vice versa. In the case of the circulant matrices we can state the following theorem, based on the numerical results in Table 4.1.

Theorem 4.5.2. *For SCTSP, the new LP relaxation (4.18) does not dominate the one tree bound, or, by implication, the Held-Karp bound (4.2).*

It was not previously known whether the SDP bound (4.1) could be worse than the 1T bound; see De Klerk, Pasechnik, and Sotirov (2008). Whether or not the Held-Karp bound dominates the new LP relaxation in the case of SCTSP remains an open question.

Theorem 4.5.3. *For SCTSP, the Held-Karp bound (4.2) is at least as tight as the Van de Veen bound (4.7).*

Proof. Let $G = (V, E)$ be a weighted circulant graph with edge weights now denoted by c_e ($e \in E$), and consider the following equivalent formulation of the Held-Karp bound (4.2) (the details may be found in Section 7.3 of Cook, Cunningham, Pulleyblank, and Schrijver (1998)):

$$
\begin{aligned}
HK = \min \quad & \sum_{e \in E} c_e x_e \\
\text{s.t.} \quad & \sum_{e \in \delta(S)} x_e \geq 2, \quad \forall\, S \subset V,\ |S| \geq 2 \\
& \sum_{e \in \delta(\{v\})} x_e = 2 \quad \forall\, v \in V \\
& 0 \leq x_e \leq 1 \quad \forall\, e \in E.
\end{aligned}
$$

We enlarge the feasible set and define a value $p^* \leq HK$ via

$$
\begin{aligned}
p^* := \min \quad & \sum_{e \in E} c_e x_e \\
\text{s.t.} \quad & \sum_{e \in \delta(S)} x_e \geq 2, \quad \forall\, S \subset V,\ S \neq \emptyset \\
& x_e \geq 0 \quad \forall\, e \in E.
\end{aligned}
$$

By LP duality theory we have

$$
\begin{aligned}
p^* = \max \quad & \sum_{\emptyset \neq S \subset V} 2 y_S \\
\text{s.t.} \quad & \sum_{S | e \in \delta(S)} y_S \leq c_e, \quad \forall\, e \in E \\
& y_S \geq 0 \quad \forall\, e \in E.
\end{aligned}
\tag{4.19}
$$

We will construct a feasible point of (4.19) with objective value equal to the value VdV from (4.7). It then follows that $p^* \geq VdV$, and since $HK \geq p^*$ we conclude that $HK \geq VdV$ for circulant matrices.

Notice that if $|V| = n$, then the dual formulation in (4.19) has $2^n - 2$ variables y_S, each corresponding to a nonempty subset of V. Let C_i^k, $(k = 0, \ldots, l-1,\ i = 1, .., \mathcal{GCD}(\Phi(k)))$, denote the connected components of the graph $G_k := \langle \Phi(1), \ldots, \Phi(k) \rangle$. In this case C_i^0 represents the vertices of the graph. According to Theorem 4.2.2, $C_i^k \neq C_j^l$ if $(i,k) \neq (j,l)$. We will abuse notation by identifying the connected component with its vertices. Define:

$$
\begin{aligned}
y_{C_i^0} &:= \frac{r_{\Phi(1)}}{2}, \quad (i = 1, \ldots, n) \\
y_{C_i^m} &:= \frac{1}{2}(r_{\Phi(m+1)} - r_{\Phi(m)}), \quad (m = 1, \ldots, l-1 \text{ and } i = 1, \ldots, \mathcal{GCD}(\Phi(m))) \\
y_S &:= 0, \ \text{otherwise.}
\end{aligned}
\tag{4.20}
$$

For a fixed m all the values $y_{C_i^m}$ are equal and nonnegative by definition, since the permutation Φ sorts r in ascending order.

According to Theorem 4.2.2 we have for each m exactly $\mathcal{GCD}(\Phi(m))$ nonzero (i.e., strictly positive) $y_{C_i^m}$ variables. Hence, the objective in (4.19) evaluates to:

$$\begin{aligned}
\sum_{\emptyset \neq S \subset V} 2y_S &= \sum_{m=0}^{l-1} 2\mathcal{GCD}(\Phi(m)) y_{C_i^m} \\
&= \mathcal{GCD}(\Phi(0)) r_{\Phi(1)} + \sum_{m=1}^{l-1} \mathcal{GCD}(\Phi(m))(r_{\Phi(m+1)} - r_{\Phi(m)}) \\
&= \sum_{m=1}^{l-1} \{(\mathcal{GCD}(\Phi(m-1)) - \mathcal{GCD}(\Phi(m))) r_{\Phi(m)}\} \\
&\quad + \mathcal{GCD}(\Phi(l-1)) r_{\Phi(l)} \\
&= \sum_{m=1}^{l} \{(\mathcal{GCD}(\Phi(m-1)) - \mathcal{GCD}(\Phi(m))) r_{\Phi(m)}\} + r_{\Phi(l)} =: VdV.
\end{aligned}$$

The last equality is due to the fact that $\mathcal{GCD}(\Phi(l)) = 1$.

To show feasibility, first fix an edge $e \in E$ with cost $r_{\Phi(k)}$, with $k \leq l$. Such an edge connects two components of G_m $(m = 0, 1, \ldots, k-1)$. Then we have

$$\begin{aligned}
\sum_{S | e \in \delta(S)} y_S &= 2\sum_{m=0}^{k-1} y_{C_i^m} = r_{\Phi(1)} + \sum_{m=1}^{k-1} (r_{\Phi(m+1)} - r_{\Phi(m)}) \\
&= r_{\Phi(1)} + r_{\Phi(k)} - r_{\Phi(1)} = r_{\Phi(k)}.
\end{aligned}$$

Now fix an edge $e \in E$ with cost $r_{\Phi(k)}$, with $k > l$. Such an edge connects at most two components of G_m $(m = 0, 1, \ldots, l-1)$. Then we have

$$\sum_{S | e \in \delta(S)} y_S \leq 2\sum_{m=0}^{l-1} y_{C_i^m} = r_{\Phi(l)} < r_{\Phi(k)}.$$

Thus, we have constructed a feasible point of (4.19) with objective value equal to the VdV bound. Therefore, HK $\geq$ VdV. $\square$

Chapter 5

Bounds for the maximum k-section problem

This chapter is based on the work of De Klerk, Pasechnik, Sotirov, and Dobre (2010). Starting from a new SDP relaxation, proposed by De Klerk and Sotirov (2010b), of the NP-complete problem of the quadratic assignment problem (QAP) we derive a new SDP bound for the maximum k-section problem, which is contained in the QAP. Further, we compare, theoretically and numerically, the resulting bounds with the existing bounds in the literature.

5.1 Introduction

Recall from Section 1.2 that the k-section (k-equipartition) of a (weighted) graph is a partition of the vertex set of the graph into k sets of equal cardinality. The weight (or cost) of a k-section is the sum of the weights of all edges that connect vertices in different sets of the partition. Thus, the maximum (resp. minimum) k-section problem is to find a k-section of maximum (resp. minimum) weight in a given weighted graph.

An equivalent formulation that will be useful is as follows. Let

$$K_{\underbrace{m,\ldots,m}_{k\text{ times}}}$$

denote a complete multipartite graph with k color classes all of size m. The maximum (resp. minimum) k-section problem is to find a $K_{m,\ldots,m}$ subgraph of maximum (resp. minimum) weight in a given weighted, complete graph on $|V| = mk$ vertices.

The maximum k-section problem is NP-hard for $k \geq 2$; see Garey, Johnson, and Stockmeyer (1976). When the weights are nonnegative, for *maximum bisection* ($k = 2$), a polynomial-time approximation ratio of 0.7016 is known from the work of Halperin and Zwick (2002) (see also Frieze and Jerrum (1997) and Ye (2001)). In other words, the randomized algorithm in Halperin and Zwick (2002) generates a bisection of the graph of expected weight at least 0.7016 times that of a maximum bisection. Andersson (1999) proposed a $(1 - 1/k + ck^3)$-approximation algorithm for maximum k-section (see also Karisch and Rendl (1998)), where c is some (unknown) absolute constant. The maximum and minimum k-section problems are different in terms of approximability (although both are NP-hard).

All the above mentioned approximation results involve SDP relaxations. In this chapter we therefore revisit SDP relaxations for max k-section and establish relationships between several SDP bounds from the literature. In particular, we present a new SDP bound for the maximum (or minimum) k-section problem, obtained from an SDP bound proposed by De Klerk and Sotirov (2010b) for the more general quadratic assignment problem. We show that the new relaxation is at least as good as the relaxation due to Frieze and Jerrum (1997) for $k = 2$ (maximum bisection). For $k \geq 3$, we prove it is at least as good as a bound introduced by Karisch and Rendl (1998). Moreover, the computation of the new SDP bound may be done much more efficiently than that of the general bound of De Klerk and Sotirov (2010b), since it requires the solution of a much smaller SDP.

This chapter is structured as follows. We first see how max k-section may be reformulated as a QAP, and then we review some known SDP relaxations of max k-section (Section 5.2). We review some SDP relaxations of QAPs in Section 5.3. These relaxations lead to large relaxations of max k-section, and to reduce the size of these SDPs we must exploit algebraic symmetry. The necessary algebraic background was presented in Section 2.2. In Section 5.3 we also derive the new SDP bound for max k-section from the QAP relaxation, by performing symmetry reduction. Theoretical comparisons with existing bounds are carried out in Section 5.4, and numerical examples are presented in Section 5.5.

5.2 Maximum k-section problem

Recall from Section 1.2.2 that the QAP reformulation of maximum k-section on a complete graph with vertex set V ($|V| = km$) and matrix of edge weights W is given by

$$\frac{1}{2}\max_{X\in\Pi_{|V|}} \operatorname{trace}(WX^TBX), \tag{5.1}$$

where B is the adjacency matrix of the complete multipartite graph $K_{m,\ldots,m}$ (m appears k times). Because of the different structures of the coherent configurations of $K_{m-1,m}$ and $K_{m-1,m,\ldots,m}$ (see Section 2.2), we treat the maximum bisection problem separately from the maximum k-section ($k > 2$) problem.

5.2.1 Maximum bisection

We are given a matrix W with nonnegative entries that we view as edge weights of a graph $G = (V, E)$, where V denotes the set of vertices and E the set of edges. In this case we consider $|V| = 2m$. The goal is to find a complete bipartite subgraph $K_{m,m}$ of G of maximum weight.

As mentioned before, maximum bisection is a special case of the more general QAP. In particular, B is the adjacency matrix of $K_{m,m}$ (with any fixed ordering of the vertices), e.g.,

$$B = \begin{pmatrix} 0_{m\times m} & J_m \\ J_m & 0_{m\times m} \end{pmatrix}. \tag{5.2}$$

An SDP relaxation of the maximum bisection problem due to Frieze and Jerrum (1997) (see also Ye (2001)) is the following:

$$\max\left\{\frac{1}{4}\operatorname{trace}(W(J_{2m}-X)) \mid \operatorname{diag}(X) = e_{2m},\ Xe_{2m} = 0,\ X \succeq 0\right\}. \tag{5.3}$$

To see that this is a relaxation of the maximum bisection problem, set $X = vv^T$, where $v \in \{-1, 1\}^{2m}$ gives the optimal equipartition of the vertex set.

5.2.2 Maximum k-section

The maximum k-section problem is a generalization of maximum bisection, where the aim is to find a complete k-equipartite subgraph of maximum total edge weight in a given weighted graph.

An SDP relaxation due to Karisch and Rendl (1998) of the max k-section problem is the following:

$$\max\left\{\frac{1}{2}\operatorname{trace}(W(J_{km}-X)) \mid \operatorname{diag}(X) = e_{km},\ Xe_{km} = me_{km},\ X \succeq 0,\ X \geq 0\right\}.$$

(5.4)

Here we present an equivalent form of the objective. If the matrix $L := \text{Diag}(Ae_n) - A$ denotes the Laplacian matrix of the graph G, then the objective of (5.4) from Karisch and Rendl (1998) is $\frac{1}{2}\text{trace}(LX)$. The equivalence can easily be seen, since:

$$\begin{aligned}
\frac{1}{2}\text{trace}(LX) &= \frac{1}{2}\text{trace}(\text{Diag}(We) - W)X \\
&= \frac{1}{2}\text{diag}(X)^T(We) - \frac{1}{2}\text{trace}(WX) \\
&= \frac{1}{2}e^TWe - \frac{1}{2}\text{trace}(WX) \\
&= \frac{1}{2}\text{trace}(W(J - X)).
\end{aligned}$$

Throughout this chapter we will refer to the SDP relaxation in (5.4) as $k - GP_{R2}$, the name given by Karisch and Rendl (1998).

Theorem 5.2.1. *If $k = 2$, the relaxation $2 - GP_{R2}$ is equivalent to the Frieze-Jerrum relaxation from (5.3).*

Proof. Given an optimal solution X of (5.3), set $\overline{X} := \frac{1}{2}(J_{2m} + X)$. Obviously, $\text{diag}(\overline{X}) = e_{2m}$ and $\overline{X} \succeq 0$. Moreover, since $Xe_{2m} = 0$ we have $\overline{X}e_{2m} = \frac{1}{2}J_{2m}e_{2m} + Xe_{2m} = me_{2m}$. Since $\text{diag}(X) = e_{2m}$ and $X \succeq 0$, its entries lie between -1 and 1, which in turn implies that the entries of $\overline{X}$ lie between 0 and 1. It is straightforward (by construction) to see that the two objective values are equal.

Conversely, assume that X is feasible for (5.4) and set $\overline{X} := 2X - J_{2m}$. We have $\text{diag}(\overline{X}) = e_{2m}$ and $\overline{X}e_{2m} = 2Xe_{2m} - J_{2m}e_{2m} = 2me_{2m} - 2me_{2m} = 0$. Since $X \succeq 0$ we have $\lambda_{\min}(X) \geq 0$. Moreover, e_{km} is an eigenvector of X with the corresponding eigenvalue equal to m. From the eigenvalue decomposition of X we have

$$X = mJ_{2m} + \sum_{i=2}^{n} \lambda_i q_i q_i^T,$$

where $\lambda_i \geq 0$ and q_i are the eigenvalues and eigenvectors of X. Then

$$X - \frac{1}{2}J_{2m} = (m - \frac{1}{2})J_{2m} + \sum_{i=2}^{n} \lambda_i q_i q_i^T,$$

and since $m - \frac{1}{2} \geq 0$ all the eigenvalues of $X - \frac{1}{2}J_{2m}$ are nonnegative. This means $X - \frac{1}{2}J_{2m} \succeq 0$, therefore $\overline{X} \succeq 0$.

It is also easy to see (by construction) that the two objectives coincide and this concludes the proof. □

5.3 New SDP bound for maximum k-section problem

The new SDP bound will be obtained by performing symmetry reduction on the SDP bound of the more general QAP. Further, we distinguish between two SDP relaxations of the QAP. One was studied by Povh and Rendl (2009), and the other by De Klerk and Sotirov (2010b).

The Povh-Rendl relaxation is as follows:

$$\left.\begin{array}{rl}
\max & \operatorname{trace}(B \otimes A)Y \\
\text{s.t.} & \operatorname{trace}(I \otimes E_{jj})Y = 1, \quad \operatorname{trace}(E_{jj} \otimes I)Y = 1 \quad (j = 1, \ldots, n) \\
& \operatorname{trace}(I \otimes (J - I) + (J - I) \otimes I)Y = 0 \\
& \operatorname{trace}(J_{n^2} Y) = n^2 \\
& Y \geq 0, \quad Y \in \mathbb{S}_+^{n^2 \times n^2},
\end{array}\right\} \tag{5.5}$$

where $I, J, E_{jj} \in \mathbb{R}^{n \times n}$. We can easily verify that (5.5) is indeed a relaxation of the QAP (1.8) by noting that a feasible point of (5.5) is given by

$$\widetilde{Y} := \operatorname{vec}(X)\operatorname{vec}(X)^T \quad \text{if } X \in \Pi_n,$$

and that the objective value of (5.5) at this point $\widetilde{Y}$ is precisely $\operatorname{trace}(BXAX^T)$.

The following discussion is condensed from De Klerk and Sotirov (2010b). Let $X \in \Pi_n$, and $r, s \in \{1, ..., n\}$ such that $X_{r,s} = 1$. Then let $\alpha = \{1, ..., n\} \setminus r$ and $\beta = \{1, ..., n\} \setminus s$. Also notice that A and B are zero diagonal, symmetric matrices. De Klerk and Sotirov (2010b) proved that the following SDP provides a lower bound for the QAP whenever the automorphism group of one of the data matrices (A or B) is transitive:

$$\left.\begin{array}{rl}
\max & \operatorname{trace}(B(\beta) \otimes A(\alpha) + \operatorname{Diag}(\bar{c}))Y \\
\text{s.t.} & \operatorname{trace}(I \otimes E_{jj})Y = 1, \quad \operatorname{trace}(E_{jj} \otimes I)Y = 1 \quad (j = 1, ..., n-1) \\
& \operatorname{trace}(I \otimes (J - I) + (J - I) \otimes I)Y = 0 \\
& \operatorname{trace}(J_{(n-1)^2} Y) = (n-1)^2 \\
& Y \geq 0, \quad Y \in \mathbb{S}_+^{(n-1)^2 \times (n-1)^2},
\end{array}\right\} \tag{5.6}$$

where $I, J, E_{jj} \in \mathbb{R}^{(n-1)\times(n-1)}$ and

$$\bar{c} := 2\operatorname{vec}(A(\alpha, \{r\})B(\{s\}, \beta)). \tag{5.7}$$

Let us now consider a matrix Y with the type of block structure that appears in (5.5) and (5.6):

$$Y := \begin{pmatrix} Y^{11} & \dots & Y^{1p} \\ \vdots & \ddots & \vdots \\ Y^{p1} & \dots & Y^{pp} \end{pmatrix}, \tag{5.8}$$

where p is a given integer and $Y^{ij} \in \mathbb{R}^{p\times p}$ for $i, j = 1, \dots, p$.

Lemma 5.3.1 (Povh and Rendl (2009)). *A matrix Y of the form (5.8) that is feasible for (5.5) (resp. (5.6)) satisfies:*

$$\operatorname{trace}(Y^{ii}) = 1 \;\; (i = 1, ..., p), \tag{5.9}$$

$$\sum_{i=1}^{p} \operatorname{diag}(Y^{ii}) = e, \tag{5.10}$$

$$e^T Y^{ij} = \operatorname{diag}(Y^{jj})^T \;\; (i, j = 1, ..., p), \tag{5.11}$$

$$\sum_{i=1}^{p} Y^{ij} = e\operatorname{diag}(Y^{jj})^T \;\; (j = 1, ..., p), \tag{5.12}$$

for $p = n$ (resp. $p = n - 1$).

In what follows, we will reduce the size of the SDP relaxation (5.6) for the QAP formulation of maximum k-section. In doing so, we will exploit the algebraic symmetry of the data matrices, i.e., the symmetry of the graph $K_{m-1,m,\dots,m}$. As in the previous section, because of the different structures of the coherent configurations of $K_{m-1,m}$ and $K_{m-1,m,\dots,m}$ (see Section 2.2), we treat the maximum bisection problem separately from the maximum k-section ($k > 2$) problem.

5.3.1 Maximum bisection

We now describe the new SDP relaxation for maximum bisection where the variables in the relaxation $X_1, \dots, X_6$ correspond to the matrices $A_1, \dots, A_6$ respectively from Example 2.2.3.

Letting $n = |V| = 2m$, $w = [W_{12} \; \dots \; W_{1n}]^T$, and

$$\bar{W} = \begin{pmatrix} W_{22} & \dots & W_{2n} \\ \vdots & & \vdots \\ W_{n2} & & W_{nn} \end{pmatrix}, \tag{5.13}$$

the new relaxation takes the following form, being obtained via symmetry reduction from (5.6):

$$\begin{aligned} SDP_{new} := \max \quad & \mathrm{trace}\,(\mathrm{diag}(w)X_5) + \frac{1}{2}\mathrm{trace}\left(\bar{W}(X_3+X_4)\right) \\ \text{s.t.} \quad & X_1 + X_5 = I_{n-1} \\ & \sum_{t=1}^{6} \mathrm{trace}(JX_t) = (n-1)^2 \\ & \mathrm{trace}(X_1) = m-1 \\ & \mathrm{trace}(X_5) = m \\ & \mathrm{trace}(X_2+X_3+X_4+X_6) = 0 \\ & X_3 = X_4^T \\ & \begin{pmatrix} \frac{1}{m-1}(X_1+X_2) & \frac{1}{\sqrt{m(m-1)}}X_3 \\ \frac{1}{\sqrt{m(m-1)}}X_4 & \frac{1}{m}(X_5+X_6) \end{pmatrix} \succeq 0 \\ & X_1 - \frac{1}{m-2}X_2 \succeq 0 \\ & X_5 - \frac{1}{m-1}X_6 \succeq 0 \\ & X_i \geq 0 \quad (i = 1, \ldots, 6). \end{aligned} \tag{5.14}$$

Note that the matrix variables X_i are all of order $n-1$.

With reference to Example 2.2.3, the reader may verify that a feasible point of the new relaxation is given by $X_i = A_i$ $(i = 1, \ldots, 6)$ if $k = m-1$ and $l = m$ in Example 2.2.3.

In what follows we show that the bound SDP_{new} in (5.14) coincides with the SDP bound for the QAP from (5.6). As mentioned before, the proof is via symmetry reduction, in the spirit of the work of Schrijver (1979) and Schrijver (2005) (see also Gatermann and Parrilo (2004)). De Klerk and Sotirov (2010b) proved that we may restrict the variable Y from (5.6) to lie in the matrix $*$-algebra

$$\mathcal{A}_{\mathrm{aut}(B(\beta))} \otimes \mathcal{A}_{\mathrm{aut}(A(\alpha))}, \tag{5.15}$$

where

$$\mathcal{A}_{\mathcal{G}} := \{X \in \mathbb{R}^{n\times n} \ : \ XP = PX, \quad \forall\, P \in \mathcal{G}\},$$

and $\mathcal{G}$ is the automorphism group of the corresponding matrix (see also Section 2.2.2). If a matrix, say B, is the adjacency matrix of a graph, then $\mathcal{A}_{\mathrm{aut}(B)}$ is a coherent algebra that contains B.

For our purpose, recall that B is the usual adjacency matrix of $K_{m,m}$, namely

$$B = \begin{pmatrix} 0 & J_m \\ J_m & 0 \end{pmatrix},$$

and $A := \frac{1}{2}W$. We fix $r = s = 1$. Hence, $\alpha = \{2, ..., n\}$, $\beta = \{2, ..., n\}$, and

$$\begin{aligned} B(\beta) &= \begin{pmatrix} 0_{m-1\times m-1} & J_{m-1\times m} \\ J_{m\times m-1} & 0_{m\times m} \end{pmatrix}, \\ A(\alpha) &= \frac{1}{2}\bar{W}, \\ \bar{c} &= \mathrm{vec}(aw^T), \end{aligned} \tag{5.16}$$

where $a^T = [0_{1\times m-1}\ e^T]$ with $e \in R^m$ the all-ones vector as before. Therefore, we can assume $Y \in \mathcal{A}_{\mathrm{aut}(K_{m-1,m})} \otimes \mathcal{A}_{\mathrm{aut}(\bar{W})}$, and since there is no symmetry assumption on the weight matrix $\bar{W}$ we have $Y \in \mathcal{A}_{\mathrm{aut}(K_{m-1,m})} \otimes \mathbb{R}^{n-1\times n-1}$.

Revisiting Example 2.2.3 we can see that $\{A_t : t = 1, ..., 6\}$ forms a basis of $\mathcal{A}_{\mathrm{aut}(K_{m-1,m})}$. Let $\{E_{ij} : i, j = 1, ..., n-1\}$ denote the standard basis of $\mathbb{R}^{n-1\times n-1}$. We can recover the basis of $\mathcal{A}_{\mathrm{aut}(K_{m-1,m})} \otimes \mathbb{R}^{n-1\times n-1}$ as $\{A_t \otimes E_{ij} : i, j = 1, ..., n-1 \text{ and } t = 1, ..., 6\}$ (for details see De Klerk and Sotirov (2010b)). Thus,

$$Y = \sum_{t=1}^{6} \sum_{i,j=1}^{n-1} y_{ij}^t A_t \otimes E_{ij},$$

for some real numbers y_{ij}^t. Further, if we denote

$$Y_t := \sum_{i,j=1}^{n-1} y_{ij}^t E_{ij}$$

we can write

$$Y = \sum_{t=1}^{6} A_t \otimes Y_t. \tag{5.17}$$

Notice that since Y is symmetric and the A_t $(t = 1, ..., 6)$ have distinct support, $Y_{t^*} = Y_t^T$ whenever $A_{t^*} = A_t^T$, for $t, t^* \in \{1, ..., 6\}$. We now substitute (5.17) in (5.6). Since the A_t are 0-1 matrices with distinct support, $Y \geq 0$ is equivalent to $Y_t \geq 0$ for $t = 1, \ldots, 6$. The positive semidefinite constraint from (5.6) becomes

$$\sum_{t=1}^{6} A_t \otimes Y_t \succeq 0. \tag{5.18}$$

If U is the unitary matrix from Theorem 2.1.5, then (5.18) is equivalent to

$$(U^* \otimes I_{n-1})(\sum_{t=1}^{6} A_t \otimes Y_t)(U \otimes I_{n-1}) \succeq 0,$$

and using (2.1) we obtain

$$\sum_{t=1}^{6} U^* A_t U \otimes Y_t \succeq 0.$$

After eliminating identical blocks from $U^* A_t U$, we reduce the matrix size of the SDP constraint in (5.6) from $(n-1)^2$ to $4(n-1)$ and write it in the form

$$\sum_{t=1}^{6} \phi(A_t) \otimes Y_t \succeq 0,$$

where ϕ is the *-isomorphism from Example 2.2.3. Defining

$$X_t := \|A_t\|^2 Y_t, \ (t = 1, ..., 6), \tag{5.19}$$

where $\|A\|$ is the Frobenius norm of matrix A, we have

$$\sum_{t=1}^{6} \frac{\phi(A_t)}{\|A_t\|^2} \otimes X_t \succeq 0.$$

Thus,

$$\frac{1}{m-1}\begin{pmatrix} X_1 & & & \\ & 0 & & \\ & & X_1 & 0 \\ & & 0 & 0 \end{pmatrix} + \frac{1}{(m-1)(m-2)}\begin{pmatrix} -X_2 & & & \\ & 0 & & \\ & & (m-2)X_2 & 0 \\ & & 0 & 0 \end{pmatrix}$$
$$+\frac{\sqrt{m(m-1)}}{m(m-1)}\begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & X_3 \\ & & 0 & 0 \end{pmatrix} + \frac{\sqrt{m(m-1)}}{m(m-1)}\begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & 0 \\ & & X_4 & 0 \end{pmatrix}$$
$$+\frac{1}{m}\begin{pmatrix} 0 & & & \\ & X_5 & & \\ & & 0 & 0 \\ & & 0 & X_5 \end{pmatrix} + \frac{1}{m(m-1)}\begin{pmatrix} 0 & & & \\ & -X_6 & & \\ & & 0 & 0 \\ & & 0 & (m-1)X_6 \end{pmatrix} \succeq 0.$$

Simplifying the last expression yields:

$$\begin{pmatrix} \frac{1}{m-1}(X_1 - \frac{1}{m-2}X_2) & & & \\ & \frac{1}{m}(X_5 - \frac{1}{m-1}X_6) & & \\ & & \frac{1}{m-1}(X_1 + X_2) & \frac{1}{\sqrt{m(m-1)}}X_3 \\ & & \frac{1}{\sqrt{m(m-1)}}X_4 & \frac{1}{m}(X_5 + X_6) \end{pmatrix} \succeq 0.$$

We now consider the linear constraints. Using (5.17), the properties of the Kronecker product (2.1) and (2.2), and the fact that only A_1 and A_5 have nonzero traces, we have

$$\begin{aligned}
\operatorname{trace}(I \otimes E_{jj})Y &= \operatorname{trace}((I \otimes E_{jj})(\sum_{t=1}^{6} A_t \otimes Y_t)) \\
&= \sum_{t=1}^{6} \operatorname{trace}(A_t \otimes E_{jj} Y_t) = \sum_{t=1}^{6} \operatorname{trace}(A_t)\operatorname{trace}(E_{jj} Y_t) \\
&= \operatorname{trace}(E_{jj}\|A_1\|^2 Y_1) + \operatorname{trace}(E_{jj}\|A_5\|^2 Y_5) \\
&= \operatorname{trace}(E_{jj}(X_1 + X_5)).
\end{aligned}$$

This yields

$$\operatorname{trace}(E_{jj}(X_1 + X_5)) = 1, \ \ (j = 1, \ldots, n-1),$$

so $X_1 + X_5 = I_{n-1}$. Continuing in the same vein,

$$\operatorname{trace}(E_{jj} \otimes I)Y = 1, \ \ (j = 1, \ldots, n-1)$$

reduces to

$$\sum_{t=1}^{6} \operatorname{trace}(E_{jj} A_t)\operatorname{trace}(Y_t) = 1, \ \ (j = 1, \ldots, n-1).$$

If we note that only $\operatorname{trace}(E_{jj}A_1)$ or $\operatorname{trace}(E_{jj}A_5)$ can be nonzero—and this can not happen for the same fixed value of j—we obtain

$$\operatorname{trace}(Y_1) = 1 \ \text{ and } \ \operatorname{trace}(Y_5) = 1.$$

Multiplying these two equations by the squared norms of A_1 and A_5 respectively we obtain two more linear equalities from (5.14), namely

$$\operatorname{trace}(X_1) = m - 1 \ \text{ and } \ \operatorname{trace}(X_5) = m.$$

Furthermore,

$$\begin{aligned}
\operatorname{trace}(\mathrm{J}_{(\mathrm{n}-1)^2}\mathrm{Y}) &= \operatorname{trace}(J_{n-1} \otimes J_{n-1})(\sum_{t=1}^{6} A_t \otimes Y_t) \\
&= \sum_{t=1}^{6} \operatorname{trace}(J_{n-1}A_t)\operatorname{trace}(J_{n-1}Y_t) = \sum_{t=1}^{6} \operatorname{trace}(J_{n-1}\|A_t\|^2 Y_t) \\
&= \sum_{t=1}^{6} \operatorname{trace}(J_{n-1}X_t).
\end{aligned}$$

This yields the following equality constraint from (5.14):

$$\sum_{t=1}^{6} \mathrm{trace}(JX_t) = (n-1)^2.$$

There is only one equality constraint left to verify. To this end let $S = \{2, 3, 4, 6\}$ and notice the following:

$$\mathrm{trace}(J-I)A_t = \begin{cases} 0 & \text{if } t \in \{1,5\} \\ \|A_t\|^2 & \text{if } t \in S. \end{cases}$$

We get

$$\begin{aligned} \mathrm{trace}((J-I)\otimes I)Y &= \mathrm{trace}((J-I)\otimes I \sum_{t=1}^{6} A_t \otimes Y_t) \\ &= \sum_{t=1}^{6} \mathrm{trace}((J-I)\otimes I)(A_t \otimes Y_t) \\ &= \sum_{t=1}^{6} \mathrm{trace}(J-I)A_t \mathrm{trace}(Y_t) = \sum_{t\in S} \mathrm{trace}(\|A_t\|^2 Y_t) \\ &= \sum_{t\in S} \mathrm{trace}(X_t). \end{aligned}$$

Also, $\mathrm{trace}(A_t) = 0$ if $t \in S$, and $X_1 + X_5 = I$, so

$$\begin{aligned} \mathrm{trace}(I\otimes(J-I))Y &= \mathrm{trace}(I\otimes(J-I) \sum_{t=1}^{6} A_t \otimes Y_t) \\ &= \sum_{t=1}^{6} \mathrm{trace}(I\otimes(J-I))(A_t \otimes Y_t) \\ &= \sum_{t=1}^{6} \mathrm{trace}(A_t)\mathrm{trace}((J-I)Y_t) \\ &= \mathrm{trace}(J-I)\|A_1\|^2 Y_1 + \mathrm{trace}(J-I)\|A_5\|^2 Y_5 \\ &= \mathrm{trace}(J-I)X_1 + \mathrm{trace}(J-I)X_5 \\ &= \mathrm{trace}(J-I)I = 0. \end{aligned}$$

We can now derive the last constraint in (5.14) immediately, since

$$\mathrm{trace}(I\otimes(J-I) + (J-I)\otimes I)Y = 0$$

is equivalent to

$$\sum_{t \in S} \operatorname{trace}(X_t) = 0.$$

The last step is to obtain the objective function. Recalling the vectors and matrices from (5.16) and equality (2.4), we can write

$$\begin{aligned}
\operatorname{trace}(\operatorname{Diag}(\bar{c}))Y &= \operatorname{trace}(\operatorname{Diag}(a) \otimes \operatorname{Diag}(w))Y \\
&= \sum_{t=1}^{6} \operatorname{trace}(\operatorname{Diag}(a)A_t)\operatorname{trace}(\operatorname{Diag}(w)Y_t) \\
&= \operatorname{trace}(\operatorname{Diag}(w)X_5),
\end{aligned}$$

where, for the last step, we used the fact that

$$\operatorname{trace}(\operatorname{Diag}(a)A_t) = \begin{cases} 0 & \text{if } t \in \{1,2,3,4,6\} \\ \|A_5\|^2 & \text{if } t = 5. \end{cases}$$

The first term of the objective function becomes

$$\begin{aligned}
\operatorname{trace}(B(\beta) \otimes A(\alpha))Y &= \frac{1}{2}\operatorname{trace}(B(\beta) \otimes \bar{W})Y \\
&= \frac{1}{2}\sum_{t=1}^{6} \operatorname{trace}(B(\beta)A_t)\operatorname{trace}(\bar{W}Y_t) \\
&= \frac{1}{2}\operatorname{trace}(\bar{W}(X_3 + X_4)),
\end{aligned}$$

where, for the last step, we used the fact that

$$\operatorname{trace}(B(\beta)A_t) = \begin{cases} 0 & \text{if } t \in \{1,2,5,6\} \\ \|A_t\|^2 & \text{if } t \in \{3,4\}. \end{cases}$$

Therefore, we have proved the following theorem.

Theorem 5.3.2. *The bound SDP_{new} from (5.14) coincides with the SDP bound (5.6) for the QAP formulation of maximum bisection.*

5.3.2 Maximum k-section

We now describe a new SDP relaxation of max k-section, $k \geq 3$, where the variables in the relaxation $X_1, \ldots, X_7$ correspond to the matrices $A_1, \ldots, A_7$ respectively in Example 2.2.4.

Letting $n = |V| = km$, the new relaxation takes the following form, obtained via symmetry reduction from (5.6):

$$
\begin{aligned}
SDP_{new} := \max \quad & \operatorname{trace}\left(\operatorname{diag}(w) X_5\right) + \frac{1}{2}\operatorname{trace}\bar{W}(X_3 + X_4 + X_7) \\
\text{s.t.} \quad & X_1 + X_5 = I_{n-1} \\
& \sum_{t=1}^{7} \operatorname{trace}(JX_t) = (n-1)^2 \\
& \operatorname{trace}(X_1) = m - 1 \\
& \operatorname{trace}(X_5) = (k-1)m \\
& \operatorname{trace}(X_2 + X_3 + X_4 + X_6 + X_7) = 0 \\
& X_3 = X_4^T \\
& \begin{pmatrix} \frac{1}{m-1}(X_1 + X_2) & \frac{1}{\sqrt{(k-1)m(m-1)}} X_3 \\ \frac{1}{\sqrt{(k-1)m(m-1)}} X_4 & \frac{1}{(k-1)m}(X_5 + X_6 + X_7) \end{pmatrix} \succeq 0 \\
& X_1 - \frac{1}{m-2} X_2 \succeq 0 \\
& X_5 - \frac{1}{m-1} X_6 \succeq 0 \\
& X_5 + X_6 - \frac{1}{k-2} X_7 \succeq 0 \\
& X_i \geq 0 \quad (i = 1, \ldots, 7).
\end{aligned}
\tag{5.20}
$$

Note that the matrix variables X_i are all of order $n-1$. With reference to Example 2.2.4, the reader may verify that a feasible point of the new relaxation is given by $X_i = A_i$ $(i = 1, \ldots, 7)$.

As mentioned before, the bound in (5.20) coincides with the SDP bound for the QAP in (5.6). The derivation is similar to the maximum bisection case, using the isomorphism in Example 2.2.4, and we therefore omit the proof and simply state the result.

Theorem 5.3.3. *For any given integer $k > 2$, the upper bound in (5.20) on the weight of a maximum k-section for a given graph coincides with the SDP bound (5.6) when applied to the QAP formulation (5.1) of maximum k-section.*

5.4 Theoretical comparison of bounds

5.4.1 Relation to the Karisch-Rendl bound

In this section we prove that the new SDP relaxation defined in (5.14) and (5.20) dominates the relaxation $k - GP_{R2}$ in (5.4), for any $k \geq 2$. The proof is slightly different for $k = 2$ and $k \geq 3$. We will present the proof only for $k \geq 3$, the proof for $k = 2$ being similar but simpler. We will need some valid implied equalities for the feasible region of (5.20). This result will follow as a consequence of Lemma 5.3.1.

Consider the following structure for the matrix variable Y of (5.6):

$$Y := \begin{pmatrix} Y^{11} & \dots & Y^{1(n-1)} \\ \vdots & \ddots & \vdots \\ Y^{(n-1)(1)} & \dots & Y^{(n-1)(n-1)} \end{pmatrix}, \tag{5.21}$$

where $Y^{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$. Following the same argument as for the case $k = 2$ and using Example 2.2.4, we can consider a variable Y of the following form:

$$Y = \sum_{t=1}^{7} A_t \otimes Y_t.$$

We then have

$$Y^{ij} = \sum_{t=1}^{7} (A_t)_{ij} Y_t, \tag{5.22}$$

where the A_t are the matrices from Example 2.2.4.

From Lemma 5.3.1

$$e^T Y^{ij} = \operatorname{diag}(Y^{jj})^T, \ \ (i, j = 1, ..., n-1). \tag{5.23}$$

Multiplying this relation by the all-ones vector to the right, we obtain

$$\operatorname{trace}(JY^{ij}) = \operatorname{trace}(Y^{jj}), \ \ (i, j = 1, ..., n-1),$$

and furthermore

$$\operatorname{trace}(JY^{ij}) = 1, \ \ (i, j = 1, ..., n-1).$$

If we substitute $i = 1$ and $j = m$ in (5.22), then $Y^{1m} = Y_3$; or if $i = m$ and $j = 1$ then $Y^{m1} = Y_4$. Continuing in the same vein, for suitable choices of i and j, we obtain

$$\operatorname{trace}(JY_t) = 1, \ \ (t = 1, ..., 7),$$

which is equivalent to

$$\text{trace}(JX_t) = \|A_t\|^2, \ \ (t = 1, ..., 7),$$

and furthermore

$$\text{trace}(JX_t) = \text{trace}(JA_t), \ \ (t = 1, ..., 7). \tag{5.24}$$

Lemma 5.4.1. *Assume the matrices $X_1, ..., X_7$ are feasible for the new SDP relaxation (5.20). Then*

$$\sum_{t=1}^{7} X_t = J, \tag{5.25}$$

$$X_1 + X_2 + X_4 = e\text{diag}(X_1)^T,$$

$$X_3 + X_5 + X_6 + X_7 = e\text{diag}(X_5)^T, \tag{5.26}$$

$$e^T X_2 = (m-2)\text{diag}(X_1)^T,$$

$$e^T X_3 = (m-1)\text{diag}(X_5)^T, \tag{5.27}$$

$$e^T X_4 = m(k-1)\text{diag}(X_1)^T, \tag{5.28}$$

$$e^T X_6 = (m-1)\text{diag}(X_5)^T,$$

$$e^T X_7 = (k-2)m\text{diag}(X_5)^T. \tag{5.29}$$

Proof. We will give the proof of (5.25), (5.26), and (5.27). The remaining equalities can be derived in a similar way.

From $\sum_{i=1}^{n-1} \text{diag}(Y^{ii}) = e$ and $\sum_{i=1}^{n-1} Y^{ij} = e\text{diag}(Y^{jj})^T$ $(j = 1, ..., n-1)$ we obtain $\sum_{i,j=1}^{n-1} Y^{ij} = J$, and further using (5.22) and the fact that the A_t, $t = 1, ..., 7$, form a coherent configuration, we get

$$\sum_{t=1}^{7} \|A_t\|^2 Y_t = J,$$

which yields $\sum_{t=1}^{7} X_t = J$, and so (5.25) is proved. To prove (5.26) we again use $\sum_{i=1}^{n-1} Y^{ij} = e\text{diag}(Y^{jj})^T$, $j = 1, ..., n-1$. If we let $j = m$ then

$$(m-1)Y_3 + Y_5 + (m-1)Y_6 + (k-2)mY_7 = e\text{diag}(Y_5)^T,$$

and using the norms of A_t, $t \in \{3, 5, 6, 7\}$,

$$X_3 + X_5 + X_6 + X_7 = e\mathrm{diag}(X_5)^T.$$

For the proof of (5.27) we use (5.22) and (5.23). If we let $i = 1$ and $j = m$,

$$e^T Y_3 = \mathrm{diag}(Y_5)^T.$$

Again using the norms of A_3 and A_5 we obtain the desired equality. □

We can now prove the main theorem of this section.

Theorem 5.4.2. *The new SDP relaxation (5.20) dominates the relaxation $k - GP_{R2}$ from (5.4).*

Proof. We will show that for any feasible point of the new SDP relaxation we can construct a feasible point of $k - GP_{R2}$ with the same objective value.

Assume that $X_1, ..., X_7$ form a feasible point for (5.20). The dimension of the all-ones vector, denoted e, can be deduced from the context and is either $n - 1$ or n.

Define

$$\widetilde{X} := \begin{pmatrix} 1 & e^T - \mathrm{diag}(X_5)^T \\ e - \mathrm{diag}(X_5) & J - (X_3 + X_4 + X_7) \end{pmatrix}. \tag{5.30}$$

The traces of X_3, X_4, and X_7 are zero so $\mathrm{diag}(\widetilde{X}) = e$. We have $X_i \geq 0$, $i = 1, ..., 7$, and $\sum_{t=1}^{7} X_t = J$ so $J - (X_3 + X_4 + X_7) \geq 0$ and further $\widetilde{X} \geq 0$.

Recall that $n = km$; using (5.27), (5.28), and (5.29) we have

$$\begin{aligned} \widetilde{X}e &= \begin{pmatrix} 1 & e^T - \mathrm{diag}(X_5)^T \\ e - \mathrm{diag}(X_5) & J - (X_3 + X_4 + X_7) \end{pmatrix} \begin{pmatrix} 1 \\ e \end{pmatrix} \\ &= \begin{pmatrix} 1 + e^T e - \mathrm{trace}(X_5) \\ e - \mathrm{diag}(X_5) + Je - (X_4^T + X_3^T + X_7^T)e \end{pmatrix} \\ &= \begin{pmatrix} 1 + (n-1) - (k-1)m \\ ne - (k-1)m(\mathrm{diag}(X_1) + \mathrm{diag}(X_5)) \end{pmatrix} \\ &= \begin{pmatrix} m \\ kme - (k-1)me \end{pmatrix} = me. \end{aligned}$$

To prove that $\widetilde{X} \succeq 0$ we use (5.26) and write $X_3 = e\mathrm{diag}(X_5)^T - (X_5 + X_6 + X_7)$. Since also $X_3 = X_4^T$ we have

$$\widetilde{X} = \begin{pmatrix} 1 & e^T - \mathrm{diag}(X_5)^T \\ e - \mathrm{diag}(X_5) & J - e\mathrm{diag}(X_5)^T - \mathrm{diag}(X_5)e^T + 2(X_5 + X_6) + X_7 \end{pmatrix}.$$

This matrix is positive semidefinite (psd) whenever the Schur complement (see Section 2.4), denoted S, of $J - e\mathrm{diag}(X_5)^T - \mathrm{diag}(X_5)e^T + 2(X_5 + X_6) + X_7$ is psd. We have

$$\begin{aligned} S &= 2(X_5 + X_6) + X_7 - \mathrm{diag}(X_5)\mathrm{diag}(X_5)^T \\ &= (X_5 + X_6) + (X_5 + X_6 + X_7) - \mathrm{diag}(X_5)\mathrm{diag}(X_5)^T. \end{aligned}$$

S is psd as the sum of two psd matrices. To see this first notice that $X_5+X_6+X_7 \succeq 0$ because

$$\begin{pmatrix} \frac{1}{m-1}(X_1 + X_2) & \frac{1}{\sqrt{(k-1)m(m-1)}}X_3 \\ \frac{1}{\sqrt{(k-1)m(m-1)}}X_4 & \frac{1}{(k-1)m}(X_5 + X_6 + X_7) \end{pmatrix} \succeq 0.$$

Further, summing $\frac{1}{k-2}(X_5 + X_6 + X_7) \succeq 0$ and $X_5 + X_6 - \frac{1}{k-2}X_7 \succeq 0$ we obtain $X_5 + X_6 \succeq 0$. Then $(X_5 + X_6 + X_7) - \mathrm{diag}(X_5)\mathrm{diag}(X_5)^T$ can be seen as the Schur complement of $X_5 + X_6 + X_7$, which is a submatrix of

$$M = \begin{pmatrix} 1 & \mathrm{diag}(X_5)^T \\ \mathrm{diag}(X_5) & X_5 + X_6 + X_7 \end{pmatrix}.$$

To conclude that $\widetilde{X} \succeq 0$ we have only to prove that $M \succeq 0$. To this end, notice that since $\mathrm{diag}(X_6) = \mathrm{diag}(X_7) = 0$ the matrix M has a special structure:

$$M = \begin{pmatrix} 1 & \mathrm{diag}(N)^T \\ \mathrm{diag}(N) & N \end{pmatrix}.$$

Using Proposition 2.4.4 we have that such a matrix is positive semidefinite if and only if $N \succeq 0$ and $\mathrm{trace}(JN) \geq \mathrm{trace}(N)^2$. We saw earlier that $N \succeq 0$; and using (5.24):

$$\begin{aligned} \mathrm{trace}(JN) &= \mathrm{trace}(J(X_5 + X_6 + X_7)) = \mathrm{trace}(J(A_5 + A_6 + A_7)) \\ &= (k-1)m + (k-1)m(m-1) + (k-1)(k-2)m^2 = (k-1)^2m^2 \\ &= \mathrm{trace}(X_5)^2 = \mathrm{trace}(N)^2. \end{aligned}$$

Therefore, $M \succeq 0$ and eventually $\widetilde{X} \succeq 0$. To conclude the proof we must show that the objective values coincide. Recall from (5.13) that

$$W = \begin{pmatrix} 0 & w^T \\ w & \overline{W} \end{pmatrix}.$$

Then

$$\begin{aligned}
\frac{1}{2}\text{trace}(W(J-\widetilde{X})) &= \frac{1}{2}w^T(e-e+\text{diag}(X_5))+\frac{1}{2}\text{trace}(w(e^T-e^T \\
&+ \text{diag}(X_5)^T)+\overline{W}(J-J+X_3+X_4+X_7)) \\
&= \frac{1}{2}(w^T\text{diag}(X_5)+\text{trace}(w\text{diag}(X_5)^T)) \\
&+ \frac{1}{2}\text{trace}(\overline{W}(X_3+X_4+X_7)) \\
&= w^T\text{diag}(X_5)+\frac{1}{2}\text{trace}(\overline{W}(X_3+X_4+X_7)) \\
&= \text{trace}(\text{diag}(w)X_5)+\frac{1}{2}\text{trace}(\overline{W}(X_3+X_4+X_7)).
\end{aligned}$$

□

Using similar techniques for $k=2$ (i.e., bisection), and defining

$$\widetilde{X} := \begin{pmatrix} 1 & e^T-\text{diag}(X_5)^T \\ e-\text{diag}(X_5) & J-(X_3+X_4) \end{pmatrix},$$

we can prove the following theorem.

Theorem 5.4.3. *The new SDP relaxation from (5.14) dominates the relaxation* $2-GP_{R2}$ *from (5.4) for any integer* $k \geq 2$.

5.4.2 Karisch-Rendl bound coincides with Povh-Rendl bound

In what follows we will show that the optimal value of the SDP relaxation $k-GP_{R2}$ coincides with the optimal value in (5.5) for the case of maximum k-section. To this end we first perform symmetry reduction of the SDP in (5.5) as in Section 5.3. As mentioned before, we can restrict the variable Y from (5.5) to lie in the matrix $*$-algebra

$$\mathcal{A}_{\text{aut}(B)} \otimes \mathcal{A}_{\text{aut}(A)}. \tag{5.31}$$

For our purpose, B is the adjacency matrix of $K_{m,\ldots,m}$ as defined in (1.10) and $A = \frac{1}{2}W$. Therefore, we can consider $Y \in \mathcal{A}_{\text{aut}(K_{m,\ldots,m})} \otimes \mathcal{A}_{\text{aut}(W)}$ and since there is no symmetry assumption on the weight matrix W we have $Y \in \mathcal{A}_{\text{aut}(K_{m,\ldots,m})} \otimes \mathbb{R}^{n\times n}$.

Revisiting Example 2.2.5 we can see that $\{A_t : t = 1,\ldots,3\}$ forms a basis of $\mathcal{A}_{\text{aut}(K_{m,\ldots,m})}$. Let $\{E_{ij} : i,j = 1,\ldots,n\}$ denote the standard basis of $n\times n$ matrices. We can choose a basis of $\mathbb{R}^{n\times n} \otimes \mathcal{A}_{\text{aut}(K_{m,\ldots,m})}$ as $\{A_t \otimes E_{ij} : i,j = 1,\ldots,n \text{ and } t = 1,\ldots,3\}$.

Then

$$Y = \sum_{t=1}^{3} \sum_{i,j=1}^{n} y_{ij}^t A_t \otimes E_{ij},$$

for some real numbers y_{ij}^t. Further, if we denote $Y_t := \sum_{i,j=1}^n y_{ij}^t E_{ij}$, we can write

$$Y = \sum_{t=1}^{3} A_t \otimes Y_t. \tag{5.32}$$

Define

$$X_t := \|A_t\|^2 Y_t,\ (t = 1, ..., 3), \tag{5.33}$$

where $\|A\|$ is the Frobenius norm of matrix A. We can now proceed by substituting (5.32) into(5.5).

Since the A_t are 0-1 matrices, $Y \geq 0$ is equivalent to $Y_t \geq 0$ for $t = 1, ..., 3$ and also equivalent to $X_t \geq 0$ for $t = 1, ..., 3$. The positive semidefinite constraint from (5.5) becomes

$$\sum_{t=1}^{3} A_t \otimes Y_t \succeq 0. \tag{5.34}$$

If I is the identity of dimension n and U is the unitary matrix from Theorem 2.1.5, then (5.34) is equivalent to

$$(U^* \otimes I)(\sum_{t=1}^{3} A_t \otimes Y_t)(U \otimes I) \succeq 0$$

and using (2.1) we obtain

$$\sum_{t=1}^{3} U^* A_t U \otimes Y_t \succeq 0.$$

Because of the commutativity of $\mathcal{A}_{\mathrm{aut}(K_{m,\ldots,m})}$, $U^* A_t U$ has a diagonal form and after deleting the repeated eigenvalues we obtain the $3n$-dimensional semidefinite constraint

$$\sum_{t=1}^{3} \phi(A_t) \otimes Y_t \succeq 0,$$

where ϕ is the *-isomorphism from Example 2.2.5. Using (5.33) we have

$$\sum_{t=1}^{3} \frac{\phi(A_t)}{\|A_t\|^2} \otimes X_t \succeq 0.$$

We can now expand the sum to obtain

$$\frac{1}{km}\begin{pmatrix} X_1 & & \\ & X_1 & \\ & & X_1 \end{pmatrix} + \frac{1}{m}\begin{pmatrix} \frac{1}{k}X_2 & & \\ & 0 & \\ & & \frac{-1}{k(k-1)}X_2 \end{pmatrix}$$
$$+\frac{1}{m}\begin{pmatrix} \frac{1}{k}X_3 & & \\ & \frac{-1}{k(m-1)}X_3 & \\ & & \frac{1}{k}X_3 \end{pmatrix} \succeq 0.$$

After simple computations we obtain the following three linear matrix inequalities:

$$\begin{aligned} X_1 + X_2 + X_3 &\succeq 0 \\ (m-1)X_1 - X_3 &\succeq 0 \\ (k-1)X_1 - X_2 + (k-1)X_3 &\succeq 0. \end{aligned} \tag{5.35}$$

Similarly to the computations carried out in Section 5.3, and using properties from (2.1) and (2.2), we obtain an equivalent formulation of the linear constraints in (5.5). Thus, $\mathrm{trace}(I \otimes E_{jj})Y = 1$, $j = 1, ..., n$, will be equivalent to $\mathrm{trace}(E_{jj}X_1) = 1$, $j = 1, ..., n$, and therefore $X_1 = I_{km}$. Also, $\mathrm{trace}(E_{jj} \otimes I)Y = 1$, $j = 1, ..., n$ will be equivalent to $\mathrm{trace}(Y_1) = 1$ and further to

$$\mathrm{trace}(X_1) = km. \tag{5.36}$$

Further, $\mathrm{trace}(I \otimes (J - I) + (J - I) \otimes I)Y = 0$ yields

$$\mathrm{trace}(X_2 + X_3) = 0. \tag{5.37}$$

Eventually, $\mathrm{trace}(JY) = n^2$ reduces to

$$\sum_{t=1}^{3} \mathrm{trace}(JX_t) = n^2. \tag{5.38}$$

The objective function becomes

$$\begin{aligned} \mathrm{trace}(B \otimes A)Y &= \frac{1}{2}\mathrm{trace}(A \otimes W)Y \\ &= \frac{1}{2}\sum_{t=1}^{3} \mathrm{trace}(AA_t)\mathrm{trace}(WY_t) \\ &= \frac{1}{2}\mathrm{trace}(WX_2). \end{aligned} \tag{5.39}$$

Therefore, the QAP relaxation of Povh and Rendl will reduce, in the case of the maximum k-section problem, to

$$\left.\begin{array}{rl} \max & \frac{1}{2}\mathrm{trace}(WX_2) \\ \text{s.t.} & \mathrm{trace}(I_{km}) + \mathrm{trace}(JX_2) + \mathrm{trace}(JX_3) = n^2 \\ & \mathrm{trace}(X_2 + X_3) = 0 \\ & I_{km} + X_2 + X_3 \succeq 0 \\ & (m-1)I_{2m} - X_3 \succeq 0 \\ & (k-1)I_{km} - X_2 + (k-1)X_3 \succeq 0 \\ & X_2, X_3 \geq 0, \end{array}\right\} \tag{5.40}$$

where the redundant constraint (5.36) has been eliminated.

To achieve our goal we need more information on the feasible set of (5.40). This extra information is obtained using linear equalities implied by (5.5), as shown in Lemma 5.3.1. We have

$$\sum_{i,j=1}^{n} Y^{(ij)} = J. \tag{5.41}$$

It follows from (5.32) that

$$Y = \begin{pmatrix} Y_1 & Y_3 & \dots & Y_3 & Y_2 & Y_2 & \dots & Y_2 \\ Y_3 & Y_1 & \dots & Y_3 & Y_2 & Y_2 & \dots & Y_2 \\ \vdots & & & \vdots & \vdots & \vdots & & \vdots \\ Y_3 & Y_3 & \dots & Y_1 & Y_2 & Y_2 & \dots & Y_2 \\ Y_2 & Y_2 & \dots & Y_2 & Y_1 & Y_3 & \dots & Y_3 \\ Y_2 & Y_2 & \dots & Y_2 & Y_3 & Y_1 & \dots & Y_3 \\ \vdots & & & \vdots & \vdots & \vdots & & \vdots \\ Y_2 & Y_2 & \dots & Y_2 & Y_3 & Y_3 & \dots & Y_1 \end{pmatrix}$$

This and (5.41) imply that

$$\|A_1\|^2 Y_1 + \|A_2\|^2 Y_2 + \|A_3\|^2 Y_3 = J$$

which further yields

$$I_{km} + X_2 + X_3 = J.$$

It is easy to see this last constraint implies the first two linear constraints and the first semidefinite constraint from (5.40), so (5.5) reduces to

$$\left.\begin{array}{rl} \max & \frac{1}{2}\mathrm{trace}(WX_2) \\ \text{s.t.} & I_{km} + X_2 + X_3 = J_{km} \\ & (m-1)I_{km} - X_3 \succeq 0 \\ & (k-1)I_{km} - X_2 + (k-1)X_3 \succeq 0 \\ & X_2, X_3 \geq 0. \end{array}\right\} \tag{5.42}$$

We can now prove the main result of this section.

Theorem 5.4.4. *The optimal values of the SDP problems (5.42) and (5.4) coincide.*

Proof. Recall that (5.4) has the following form ($n = mk$):

$$\max\left\{\frac{1}{2}\mathrm{trace}(W(J_{km} - X)) \mid \mathrm{diag}(X) = e_{km},\ Xe_{km} = me,\ X \succeq 0,\ X \geq 0\right\}.$$

Given an optimal solution X_2, X_3 of (5.42), set

$$X := J_{km} - X_2 \succeq 0.$$

The SDP (5.42) appears as problem (10) in Section 4.3 of the paper by De Klerk and Pasechnik (2009). It therefore follows from Theorem 3.1 in Section 3 of that paper that $X_2 e_{km} = m(k-1)e_{km}$. Hence,

$$\begin{aligned} Xe_{km} &= J_{km}e_{km} - X_2 e_{km} \\ &= mke_{km} - m(k-1)e_{km} = me_{km}, \end{aligned}$$

where we have used the fact that $A_2 e_{km} = m(k-1)e_{km}$, for the association scheme of $K_{m,\ldots,m}$ (see Example 2.2.5). Moreover, it is easy to verify that $\mathrm{diag}(X_2) = 0$, so $\mathrm{diag}(X) = e_{km}$. We have $X = I_{km} + X_3 \geq 0$; and obviously the objective values of the two problems coincide.

Conversely, assume that X is feasible for (5.4). Setting

$$X_2 = J_{km} - X,\ \ X_3 = X - I_{km}$$

yields a feasible solution of (5.42) with the same objective function. We have $\mathrm{diag}(X) = e_{km}$ and $X \succeq 0$ so $X_{ij} \in (-1, 1)$ and $X_2 \geq 0$. Also, $\mathrm{diag}(X) = e_{km}$ and $X \geq 0$ so $X_3 \geq 0$ and obviously $I_{km} + X_2 + X_3 = J_{km}$.

The LMI $(k-1)I_{km} - X_2 + (k-1)X_3 \succeq 0$ is equivalent to $X - \frac{1}{k}J \succeq 0$. To see that the latter is true, notice that e_{km} is an eigenvector of X and the only eigenvalue that changes is its corresponding eigenvalue. Moreover, this eigenvalue stays nonnegative since $e_{km}^T(X - \frac{1}{k}J)e_{km} = e_{km}^T m e_{km} - \frac{1}{k}(mk)^2 = 0$.

The LMI $(m-1)I_{km} - X_3 \succeq 0$ is equivalent to $mI_{mk} - X \succeq 0$. The spectral radius of X, $\rho(X) \leq \|X\|_\infty$ and in our case $\|X\|_\infty = m$, so no eigenvalue is larger than m. $\square$

5.5 Numerical results

In this section we present numerical results for the new SDP bound (5.14) and (5.20), the bound due to Frieze and Jerrum (5.3), and the bound due to Karish and Rendl, $k - GP_{R2}$. The matrices have dimensions between 9 and 30 in order to be tractable with all the approaches. The column "*times*" in the tables gives the times in seconds to compute the new SDP relaxation on a dual core Pentium IV (2×2, 13 GHz) with 2 GB of RAM. The times reported for the new SDP bound include the time to solve n SDP relaxations, corresponding to n distinct fixings of rows and columns.

In the first table we deal with minimization (to compare with existing results for minimum bisection), and the second table presents computational results and times for maximum 3-equipartition.

The instances denoted by R and a number are randomly generated, up to dimension 21, so that we could also solve them to optimality by exact enumeration. The instances *cb*.30.47 and *cb*.30.56 were taken from the PhD thesis of Ambruster (2007). The optimal values of these problems were reported in Table C.50 of Appendix A on page 203 of that thesis.

The instances from Table 5.1 and Table 5.2 are available online at:
`http://lyrawww.uvt.nl/~cdobre/equipart_instances.rar`.

Table 5.1: Bounds on optimal values of min bisection

problem	dimension	time (s)	FJ	new SDP	optimum
R1	14	88	4,316,3	4,375.1	4,387
R2	12	33	3,267.9	3,300	3,300
R3	16	185	531.4	538	538
R4	18	356	694.6	701.9	709
R5	20	715	767.3	773	773
cb.30.47	30	10,447	201.22	213	266
cb.30.56	30	10,139	291.82	302	379

Table 5.2: Bounds on optimal values of max 3-equipartition

problem	dimension	time (s)	3GPR2	new SDP	optimum
R6	9	5.47	2,774.54	2,773	2,773
R7	12	39.28	5,265.58	5,255	5,255
R8	15	179.37	8,095.34	8,029.87	8,000
R9	18	676.49	11,526.20	11,460.04	11,459
R10	21	1,743.1	16,316.74	16,238.74	16,175

Bibliography

Alizadeh, F. (1991). *Combinatorial Optimization with Interior Point Methods and Semidefinite Matrices*. Ph. D. thesis, University of Minnesota, Minneapolis, USA.

Ambruster, M. (2007). *Branch-and-Cut for a Semidefinite Relaxation of Lagrange-Scale Minimum Bisection Problem*. Ph. D. thesis, Chemnitz University of Technology, Germany.

Andersson, G. (1999). An approximation algorithm for max p-section. In C. Meinel and S. Tison (Eds.), *STACS'99*, pp. 237–247. Singapore: LNCS 1563.

Anstreicher, K.M. (2003). Recent advances in the solution of quadratic assignment problems. *Mathematical Programming B*, **97**, 27–42.

Babel, L., S. Baumann, M. Lüdecke, and G. Tinhofer (1997). STABCOL: Graph isomorphism testing based on the Weisfeiler-Leman algorithm. Report TUM M9702, Mathematisches Institut, Technische Universität München.

Bachoc, C. and F. Vallentin (2008). New upper bounds for kissing numbers from semidefinite programming. *Journal of the American Mathematical Society*, **21**, 909–924.

Bai, Y. Q., E. de Klerk, D. V. Pasechnik, and R. Sotirov (2009). Exploiting group symmetry in truss topology optimization. *Optimization and Engineering*, **10(3)**, 331–349.

Bogdanowicz, Z. (2005). Hamiltonian circuits in sparse circulant digraphs. *Ars Combinatoria*, **76**, 213–223.

Boyd, S. and L. Vandenberghe (2004). *Convex Optimization*. Cambridge University Press.

Burkard, R.E. and W. Sandholzer (1991). Efficiently solvable special cases of the bottleneck traveling salesman problem. *Discrete Applied Mathematics*, **32**, 61–76.

Cameron, P.J. (2003). Coherent configurations, association schemes and permutation groups. In A.A. Ivanov, M.W. Liebeck, and J. Saxl (Eds.), *Groups, Combinatorics and Geometry*, pp. 55–71. Singapore: World Scientific.

Codenotti, B., I. Gerace, and S. Vigna (1988). Hardness results and spectral techniques for combinatorial problems on circulant graphs. *Linear Algebra and its Applications*, **285**, 123–142.

Cook, W.J., W.H. Cunningham, W.R. Pulleyblank, and A. Schrijver (1998). *Combinatorial Optimization.* New York, USA: John Wiley and Sons.

Dantzig, G.B., D.R. Fulkerson, and S.M. Johnson (1954). Solution of a large scale traveling salesman problem. *Operations Research*, **2**, 393–410.

Dukanovic, I. and F. Rendl (2007). Semidefinite programming relaxations for graph coloring and maximal clique problems. *Mathematical Programming B*, **109**, 345–365.

Eberly, W. and M. Giesbrecht (2004). Efficient decomposition of separable algebras. *Journal of Symbolic Computation*, **37(1)**, 35–81.

Etingof, P., O. Golberg, S. Hensel, T. Liu, A. Schwendner, E. Udovina, and D. Vaintrob (2009). Introduction to representation theory. arXiv:0901.0827v3.

Frieze, A. and M. Jerrum (1997). Improved approximation algorithms for MAX k-CUT and MAX BISECTION. *Algorithmica*, **18(1)**, 67–81.

GAP (2008). Groups, algorithms, and programming, Version 4.4.12. The GAP Group, `http://www.gap-system.org`.

Garey, M.R., D.S. Johnson, and L. Stockmeyer (1976). Some simplified NP-complete graph problems. *Theoretical Computer Science*, **3(1)**, 237–267.

Garfinkel, R.S. (1977). Minimizing wallpaper waste, Part 1: A class of TSP. *Operations Research*, **25**, 741–751.

Gatermann, K. and P.A. Parrilo (2004). Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, **192**, 95–128.

Gerace, I. and F. Greco (2008). The symmetric circulant traveling salesman problem. In F. Greco (Ed.), *Travelling Salesman Problem.* Wien: I-Tech.

Gijswijt, D. (2005). *Matrix Algebras and Semidefinite Programming Techniques for Codes*. Ph. D. thesis, University of Amsterdam, The Netherlands.

Gijswijt, D., A. Schrijver, and H. Tanaka (2006). New upper bounds for nonbinary codes based on the Terwilliger algebra and semidefinite programming. *Journal of Combinatorial Theory A*, **113**, 1719–1731.

Godsil, C. (2005). *Association Schemes*. Lecture notes, University of Waterloo, Available at: `http://quoll.uwaterloo.ca/mine/Notes/assoc.pdf`.

Goemans, M.X. and D.P. Williamson (1995). Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the Association for Computing Machinery*, **42(6)**, 1115–1145.

Graham, A. (1981). *Kroneker Products and Matrix Calculus with Applications*. Ellis Horwood Limited.

Gray, R.M. (2006). Toeplitz and circulant matrices: A review. *Foundation and Trends in Comunications and Information Theory*, **2**, 155–239.

Gvozdenović, N. and M. Laurent (2008a). Computing semidefinite programming lower bounds for the (fractional) chromatic number via block-diagonalization. *SIAM Journal on Optimization*, **19(2)**, 592–615.

Gvozdenović, N. and M. Laurent (2008b). The operator Ψ for the chromatic number of a graph. *SIAM Journal on Optimization*, **19(2)**, 572–591.

Halperin, E. and U. Zwick (2002). A unified framework for obtaining improved approximation algorithms for maximum graph bisection problems. *Random Structures and Algorithms*, **20(3)**, 382–402.

Held, M. and R.M. Karp (1970). The traveling salesman problem and minimum spanning trees. *Operations Research*, **18**, 1138–1162.

Heuberger, C. (2002). On Hamiltonian Toeplitz graphs. *Discrete Mathematics*, **245**, 107–125.

Higman, D.G. (1987). Coherent algebras. *Linear Algebra Appl.*, **93**, 209–239.

Horn, R.A. and C.R. Johnson (1990). *Matrix Analysis*. Cambridge, MA, USA: Cambridge University Press.

Kanno, Y., M. Ohsaki, K. Murota, and N. Katoh (2001). Group symmetry in interior-point methods for semidefinite program. *Optimization and Engineering*, **2(3)**, 293–320.

Karisch, S.E. and F. Rendl (1998). Semidefinite programming and graph equipartition. In P.M. Pardalos and H. Wolkowicz (Eds.), *Topics in Semidefinite and Interior-Point Methods*. American Mathematical Society.

Klerk, E. de (2010). Exploiting special structure in semidefinite programming: A survey of theory and applications. *European Journal of Operational Research*, **201(1)**, 1–10.

Klerk, E. de and C. Dobre (2009). A comparison of lower bounds for the symmetric circulant traveling salesman problem. Preprint, Tilburg University.

Klerk, E. de, C. Dobre, and D.V. Pasechnik (2010). Numerical block diagonalization of matrix $*$-algebras with application to semidefinite programming. *Mathematical Programming B. To appear*.

Klerk, E. de, J. Maharry, D.V. Pasechnik, B. Richter, and G. Salazar (2006). Improved bounds for the crossing numbers of $K_{m,n}$ and K_n. *SIAM Journal on Discrete Mathematics*, **20**, 189–202.

Klerk, E. de, M.W. Newman, D.V. Pasechnik, and R. Sotirov (2009). On the Lovász ϑ-number of almost regular graphs with application to Erdös-Rényi graphs. *European Journal of Combinatorics*, **31**, 879–888.

Klerk, E. de and D.V. Pasechnik (2002). Approximation of the stability number of a graph via copositive programming. *SIAM Journal on Optimization*, **12(4)**, 875–892.

Klerk, E. de and D.V. Pasechnik (2009). On semidefinite programming relaxations of association schemes with applications to combinatorial optimization problems. Discussion paper no 2009-54, CentER.

Klerk, E. de, D.V. Pasechnik, and A. Schrijver (2007). Reduction of symmetric semidefinite programs using the regular *-representation. *Mathematical Programming B*, **109(2-3)**, 613–624.

Klerk, E. de, D.V. Pasechnik, and R. Sotirov (2008). On semidefinite programming relaxations of the traveling salesman problem. *SIAM Journal on Optimization*, **19(4)**, 1559–1573.

Klerk, E. de, D.V. Pasechnik, R. Sotirov, and C. Dobre (2010). On semidefinite programming relaxations of maximum k-section. Preprint, Tilburg University.

Klerk, E. de and R. Sotirov (2010a). Exploiting group symmetry in semidefinite programming relaxations of the quadratic assignment problem. *Mathematical Programming A*, **122(2)**, 225–246.

Klerk, E. de and R. Sotirov (2010b). Improved semidefinite programming bounds for quadratic assignment problems with suitable symmetry. *Mathematical Programming A. To appear*.

Koopmans, T.C. and M.J. Beckmann (1957). Assignment problems and the location of economics activities. *Econometrica*, **25**, 53–76.

Korte, B. and J. Vygen (2008). *Combinatorial Optimization: Theory and Algorithms, 4th edition. Algorithms and Combinatorics 21.* Berlin: Springer-Verlag.

Laurent, M. (2009). Strengthened semidefinite bounds for codes. *Mathematical Programming B*, **109(2-3)**, 239–261.

Laurent, M., S. Poljak, and F. Rendl (1997). Connections between semidefinite relaxations of the max-cut and stable set problems. *Mathematical Programming B*, **77(2)**, 225–246.

Löfberg, J. (2004). YALMIP: A toolbox for modeling and optimization in MATLAB. Proceedings of the CACSD conference, Taipei, Taiwan, `http://control.ee.ethz.ch/~joloef/yalmip.php`.

Lovász, L. (1979). On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, **25**, 1–7.

McEliece, R.J., E.R. Rodemich, and H.C. Rumsey (1978). The Lovász bound and some generalizations. *Journal of Combinatorics, Information and System Sciences*, **3**, 134–152.

Medova, E. (1994). Using QAP bounds for the circulant TSP to design reconfigurable networks. *Discrete Mathematics and Theoretical Computer Science*, **16**, 275–292.

Mittelmann, H.D. and F. Vallentin (2010). High accuracy semidefinite programming bounds for kissing numbers. *Experimental Mathematics*, **19**, 174–178.

Murota, K., Y. Kanno, M. Kojima, and S. Kojima (2010). A numerical algorithm for block-diagonal decomposition of matrix *-algebras with application to semidefinite programming. *Japan Journal of Industrial and Applied Mathematics*, **27(1)**, 125–160.

Nesterov, Yu.E. and A.S. Nemirovski (1994). *Interior Point Polynomial Algorithms in Convex Programming.* SIAM Studies in Applied Mathematics, Vol. 13. Philadelphia, USA: SIAM.

Povh, J. and F. Rendl (2009). Copositive and semidefinite relaxations of the quadratic assignment problem. *Discrete Optimization*, **6(3)**, 231–241.

Rendl, F. and R. Sotirov (2007). Bounds for the quadratic assignment problem using the bundle method. *Mathematical Programming B*, **109(2-3)**, 505–524.

Schrijver, A. (1979). A comparison of the Delsarte and Lovász bounds. *IEEE Transactions on Information Theory*, **25**, 425–429.

Schrijver, A. (2003). *Combinatorial Optimization, Volume B - Polyhedra and Efficiency.* Springer.

Schrijver, A. (2005). New code upper bounds from the Terwilliger algebra. *IEEE Transactions on Information Theory*, **51**, 2859–2866.

Shor, N.Z. (1987). Quadratic optimization problems. *Soviet Journal of Computer and System Sciences*, **25**, 1–11.

Soicher, L.H. (2006). GRAPE: GRaph Algorithms using PErmutation groups. A GAP package, Version number 4.3, `http://www.maths.qmul.ac.uk/~leonard/grape/`.

Sturm, J.F. (1999). Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, **11/12**, 625–653.

Todd, M.J. (2001). Semidefinite optimization. *Acta Numerica*, **10**, 515–560.

Toh, K.C., M.J. Todd, and R.H. Tütüncü (1999). SDPT3—a MATLAB software package for semidefinite programming, version 1.3. *Optimization Methods and Software*, **11/12(1-4)**, 545–581.

Turán, P. (1977). A note of welcome. *Journal of Graph Theory*, **1**, 7–9.

Vallentin, F. (2009). Symmetry in semidefinite programs. *Linear Algebra and Applications*, **430**, 360–369.

Vandenberghe, L. and S. Boyd (1996). Semidefinite programming. *SIAM Review*, **38**, 49–95.

Veen, J.A.A. van der (1992). *Solvable Cases of TSP with Various Objective Functions.* Ph. D. thesis, Groningen University, The Netherlands.

Wedderburn, J.H.M. (1907). On hypercomplex numbers. *Proceedings of the London Mathematical Society*, **6(2)**, 77–118.

Yang, Q.F., R.E. Burkard, E. Cela, and G.J. Wöginger (1997). Hamiltonian cycles in circulant digraphs with two stripes. *Discrete Mathematics*, **176**, 233–254.

Ye, Y. (2001). A .699 approximation algorithm for max-bisection. *Mathematical Programming, Series A*, **90**, 101–111.

Zhao, Q., S.E. Karisch, F. Rendl, and H. Wolkowicz (1998). Semidefinite programming relaxations for the quadratic assignment problem. *Journal of Combinatorial Optimization*, **2**, 71–109.